

Are federations always indissoluble?

Russia's heavy-handed approach to separatist tendencies in Chechenia is old-fashioned, to say the least, especially because both federalism and separatism are likely to be prominent features of the constitutional landscape ahead.

BOTH the United States and what is now loosely called Russia are federations in the constitutional lawyers' sense. A little over 130 years ago, the United States fought one of the most destructive wars in its history to prevent its southern states seceding from the Union established less than a century earlier. Now Mr Boris Yeltsin, the president of the Russian Federation, is engaged on a similar struggle to prevent the southern state of Chechenia from gaining the independence to which it has asserted it has a right. In many ways, Yeltsin's war has begun as badly as that of Abraham Lincoln. The constitutional forces seeking to enforce the integrity of the federation have been taught many nasty lessons by the rebels, there has been sad (and needless) loss of life and wanton destruction of property in a region of Russia not well endowed with material riches. But Yeltsin, if he wishes, will no doubt eventually have his way. Does not the precedent of the American Civil War entitle him to no less?

That would be a false claim if it were ever made. The United States of America was a voluntary arrangement between previously equipotent entities which had in common their subservience to the British Crown as colonies. They would not have rid themselves of that dependence had they not made common cause and mounted what British history-books still call the 'American Revolution'. The case against the secession of the South is that it would not have had something from which to secede were it not for the federation, but also that, after several decades of mutual coexistence and economic evolution, the North would have had as much to lose from disunion as the South could possibly have hoped to gain. (The abolition of slavery became an issue only after the Civil War had begun.)

The Russian Federation does not have as strong a case to keep Chechenia in the fold. Although the occupation of the southern republics predates the formation of the Soviet Union, and despite the uniformity enforced in the past 70 years, there has not been an occasion when Chechens can be said voluntarily to have thrown in their cause with the government of the day in Moscow. Constitutional lawyers would, no doubt, be quick to say that the manner of Chechenia's incorporation in the Russian Federation was legitimate by the standards applicable at the time, and so is now also legitimate. But that is an old-fashioned argument.

To say the least, that consideration should have moderated the severity of Moscow's reaction in the past few weeks and should have prompted negotiation rather than force. In the event, Yeltsin may have done himself more harm than

good. Statesmen these days win very few friends by sending in the troops. But Yeltsin will also have incurred a debt to the military that he may be hard pressed to repay. And by making Chechenia such a conspicuous issue, he will have dramatized secession in the minds of those whose separation from the Russian Federation would be much more serious than that of the Chechens. Will the peoples of Russia's Far East now rediscover separatist hankerings?

That implies that there are occasions when federations can legitimately be dissolved. It is only a few years since, for example, the Swiss canton of the Jura separated from the Swiss Confederation and became part of France. (The cantons of the Swiss Confederation differ from the states of a federal state in ceding to the federal government only functions such as foreign policy and defence.) The principle is that the dissolution of a long-standing arrangement requires the consent both of those who wish to separate and those who will be left behind. (That condition did not obtain in the United States 135 years ago, whence the Civil War.) In modern circumstances, it does not suffice that the government of the day should give or withhold its consent. A more general expression of public opinion is required. Yeltsin has not sought it in respect of Chechenia.

That omission apart, there will be a growing tendency for the components of federal states to shake themselves loose from the larger entities to which they belong, especially in Europe, where the European Union has strengthened regional ambitions to be distinctive. The identity of the German *Länder* is constitutional as well as traditional, but the recent strengthening of regional identity in France is more remarkable. Belgium has gone even further, shaking down into two linguistic areas. Whether Italy follows suit remains to be seen. In the circumstances, it is unremarkable that even in Britain, Scotland's second-largest political party advocates independence from the United Kingdom.

How can these issues be resolved? In the case of Scotland, the present British government considers even the further devolution of power to Scotland to be anathema, but its principal political opponent (the Labour Party) offers devolution as a consequence of its success at the next general election. The issue has the makings of a fierce political argument, but a meaningless one. Britain already administers Scotland separately, through the Scottish Office. Adding a degree of regional political representation would be only sensible. The consent of both the separatists and those left behind would be readily won if the question were put



fairly, especially because common membership of the European Union ensures that people can move freely from one place to another.

The more taxing question is that of Britain's membership of the European Union itself. That is already at least a confederation of states on the Swiss pattern. Although member states still retain control of foreign policy and defence, self-interest dictates that they work increasingly in concert. Meanwhile, the volume of legislation promulgated from the centre far exceeds what Swiss cantons would stomach. Although any member could probably still withdraw from the union without suffering serious penalties, that will not indefinitely be the case. Yet the British government continues to insist that federalism is not on the European agenda. Constitutionally, it may not be, but in reality it is unavoidable. That is why clever European states are busily staking out boundaries of their cultural identity that may survive the processes under way. □

Complexity on a chip

Do not pity those compelled to do without their latest microprocessor.

INTEL, the US manufacturer of microprocessors mounted on silicon chips, has sadly dented its own reputation for competence in the past few months, and needlessly. The tale has now been told countless times. Intel's latest and much advertised microprocessor, called "Pentium", was found to be defective in design as long ago as last September. The built-in co-processor (used, among other things, for floating point calculations) was discovered not always to perform accurately. Unfortunately, the company did not make a clean breast of what it knew when the error first came to light, but rather let its competitors draw attention to it at the season in the consumer year when fond parents are most inclined to buy expensive presents for their offspring, in the run-up to the Christmas holiday. Nobody can accurately estimate how many computers were not sold because of the bad publicity.

This is not, of course, the only occasion when a manufactured chip has failed to perform as the designers intended. Moreover, it is inevitable that the increasing complexity of the circuits now etched on silicon must increase the difficulty of detecting errors in design or manufacture. In a closely related field, it is well-known that complex software programs are unlikely ever to be tested in all the circumstances in which they are meant to function. In that spirit, chip-manufacturers should be ready to acknowledge that their craft has reached the point of complexity at which simulated advance testing cannot be entirely effective, and should be ready to share their anxiety with their customers. What that implies is what manufacturers of other products, aircraft for example, have long been familiar with: the time elapsed between the conception of a new design and its successful sale is prolonged by seemingly endless proving trials. But time means money, does it not? Of course. But not endless money. And who will seriously weep for the army of

computer users who must rub along with Pentium's predecessor, the "486", while the testing is carried out? □

Darwin lives on . . .

A British magazine appears to believe (with Lysenko) that genes can triumph over natural selection.

WILL Darwinism follow Marxism into oblivion? That is the question raised by *The Spectator*, a largely political British weekly magazine even older than *Nature*, whose origins (in 1828) lie in the social ferment preceding the great reform bills of 1830 and 1832. More recently, *The Spectator* has forsaken its liberal beginnings for what may easily be mistaken for their opposites. Now, after many months of exultation that victory in the Cold War has gone to the like-mindedly righteous, it has given houseroom to four pages of speculation on what will happen now that Darwinism is dead. The author is one Warwick Collins, whose claim on public attention appears to be that he was once a student of Professor John Maynard Smith.

The interest of Collins's article is not that it should have been written — people are forever writing all kinds of things — but that it should have been published so portentously. *The Spectator* says, "now that the environmental theories of Marxism have collapsed..., perhaps the same fate will befall Darwinism". That, in itself, is a very curious proposition. Neither the magazine nor Collins appear to have remembered that Soviet-style Marxism backed Lysenko against Darwin precisely because it could not stomach the idea that people cannot triumph over their genomes (or those of cereals) by taking sufficient thought.

Sadly, there is worse to come. Thus Collins writes that the "premise [of Darwinism] that the organism becomes adapted to its environment, and is thus 'designed' by the environment, appears absurd. Precisely how may a passive environment design a highly active stream of organisms?" No wonder, it may be thought, that Collins appears not to have completed his time as a student of Maynard Smith. The premise of Darwinism is that there is not adaptation but variation, and that less well-adapted variants are less fit (in the technical sense of being less likely to perpetuate their genes). Beginning students fail examinations for making such mistakes although Collins, undeterred, goes on to argue that Darwinism entails the phenomenon of "pre-adaptation" (clairvoyance by an organism's genome) and that natural selection suppresses variation.

What is all this in aid of? "To our modern eyes, ... the vital processes of evolution are powered ... by indigenous processes within organisms themselves." Shades of Lamarck? Collins has something more ambitious in mind. "Darwin's view that evolving organisms are the product of the environment is likely to be superseded by the belief that it is the active indigenous processes of those organisms which shape and determine that environment." Gaia made conscious and Lysenko made respectable. Is that what *The Spectator* intended by giving space to Collins? □

Republicans see private sector as holding key to growth of US science

Washington. Tax incentives and deregulation to encourage more research in the private sector are the best means to expand US science in the future, Robert Walker (Republican, Pennsylvania), new chairman of the House of Representatives Science Committee, said last week.

Walker was speaking after a packed hearing on Capitol Hill to which he had summoned the heads of the main federal science agencies to provide their respective visions for the next 20 years.

"I want to see resources for science grow in the total economy," said Walker. "That doesn't mean bigger government programmes." He promised to use his strong ties with the rest of the new Republican leadership in the House to win tax concessions for programmes he favours, such as private-sector space launches and the development of hydrogen-powered cars.

The hearing was the first chance for the agency chiefs of the Clinton administration to exchange ideas (or cross swords) with the new Republican Congress. Those present included Jack Gibbons, the President's Science Advisor; Dan Goldin, administrator of the National Aeronautics and Space Administration (NASA); Neal Lane, director of the National Science Foundation (NSF); Carol Browner, administrator of the Environmental Protection Agency and Ron Brown, the Secretary of Commerce.

Walker's committee has jurisdiction over most federal science and technology (with the important exceptions of biomedical and defence research). The pugnacious Pennsylvania congressman has also been asked by his friend Newt Gingrich, the Speaker of the House, to serve as vice-chair on the powerful Budget Committee, where he will exert strong influence on federal spending.

But the hearing suggested that, although money will be tight, Republicans on the Science Committee at least have no intention of allowing an all-out attack by fiscal

conservatives on the \$73-billion federal science and technology budget.

Gibbons may have captured the administration's expectations adroitly by quoting Samuel Johnson's phrase that when a man knows he is to be hanged, "it concentrates his mind wonderfully". But he must have been quietly impressed by the number of Republican congressmen whose questions implied that they had turned up chiefly to defend science and technology programmes in their districts.

Before an unprecedentedly large crowd that packed the hall for three hours while many more were shut outside, the testimony of the agency chiefs revealed or confirmed a number of important specifics. Dan Goldin, for example, while declining to discuss the pending administration budget proposal for the 1996 financial year, made it clear there will be no money yet for the construction of a \$2.5-billion wind-tunnel complex (see *Nature* 370, 242; 1994). That will mean the automatic rescission of \$400 million appropriated for the complex by Congress last year.

Alone among the agency chiefs — and true to form — Goldin accepted the committee's invitation to look at the year 2015 through a crystal ball. He foresaw robots on the Moon operated by schoolchildren and commercial airliners flying at Mach 2.5 (2.5 times the speed of sound), generating \$250 billion in sales.

Neal Lane of the NSF declined to follow suit. He did not have "the slightest idea" what 2015 will look like, he said. Such predictions were "notoriously wrong". he

most as good as new", according to ILL.

More than 1,400 researchers are expected to use the ILL this year. Almost 600 requests for experiments have been registered for the first semester alone, representing more than 2.5 times the amount of beamtime available.

But the number of beamlines has been reduced from 31 to 25 because of budgetary restrictions imposed by the United Kingdom, which, along with France and Germany, is one of the main contributors to the ILL.

D.B.

added. "Proud. But wrong."

Ron Brown was unable to convince Walker that his department's technology

programmes are worth saving. They are "a critical fraction of a percent" of federal research and development, he said, and to lose them "would be an act of unilateral disarmament by the United States". But Walker said afterwards that the largest of them — the Advanced Technology Program — should be drastically curtailed.

Browner, whose Environmental Protection Agency is on course for a direct confrontation with the new Congress on several fronts, bluntly spelt out her differences with the Republicans on the issue of risk assessment. Under the proposals for this in the Republican manifesto, the *Contract with America*, she said, "we could not have banned lead in gasoline".

After the hearing, Gibbons said that he did not in fact expect US science to be hanged by the 104th Congress. "We are entering a period of focused discussion on these issues," he said. "The more we can talk about them, the more we will raise understanding of the fundamental value of science." Walker naturally concurred. "In my time in the House [there has never been] a leadership more enthusiastic about science and technology", he said.

It took Vernon Ehlers (Republican, Michigan) — until recently an active physicist — to raise scientists' concern about the intentions of both Republicans and Democrats. "I don't see the nation having the commitment to basic research that we had before. I see it faltering," he said.

George Brown (Democrat, California) and Tim Roemer (Democrat, Indiana) disputed the main thrust of the testimony from Gibbons, Goldin and Lane, — that in future they will be doing more science for less money. "Dan Goldin says we're going to see children playing with robots on the moon, but I just don't see it," said Roemer, an erstwhile science supporter and space station opponent. "We are going to have to make some tough cuts in these programmes. What about reality? How are you going to make these cuts?"

"You get efficient, you rationalize, and you measure outputs, not inputs," replied Goldin. "There's lots of room to eliminate lots of things we do." **Colin Macilwain**

ILL reopens after 4-year shut-down

Paris. The world's most powerful neutron source, the 58.3 MW reactor at the Institut Laue-Langevin (ILL) in Grenoble, France, started up again last week after having been closed down for repairs for almost four years.

The reactor was shut down in 1991 after 20 years of operation, following the discovery of cracks inside the apparatus. The subsequent repairs, which cost FF173 million (US\$32.2 million) and included replacing the entire reactor vessel, mean that the refurbished reactor is "al-

Call for more coordination of Gulf war syndrome research

Washington. Efforts to identify the so-called 'Gulf War Syndrome' affecting US troops who served in the Gulf War largely consist of badly designed and poorly coordinated studies that are unlikely to provide useful results, according to the US Institute of Medicine (IoM).

An IoM panel describes the 50 or so separate public and private studies now under way as well-intentioned. But it claims that they are "intrinsically unsuitable for systematic study of the health effects of the Gulf War", because of the methods used to select participants and collect data.

The IoM panel, chaired by John Bailer of McGill University, Montreal, Canada, found that most of the studies were piecemeal in nature, and that some had "several fundamental problems". Most lacked proper controls to compare, for example, the health of the Gulf veterans being studied with a control group of other subjects.

"Overall, there has been a serious lack of adequate attention to the design of individual studies," the panel found, while the failure to coordinate the studies was

an "even more serious" weakness.

Part of the problem has been a lack of coordination between the main government departments involved, namely Defense, Veterans' Affairs, and Health and Human Services. The IoM calls for Vice-President Al Gore to chair a committee to coordinate their efforts, and to involve leading epidemiologists.

But Bailer says that Defense and Veterans' Affairs officials who saw the report before publication appeared receptive to its central message. "What's been done is characteristic of the early stages of recognition of any new or newly recognized disease," he says. That phase has now come to an end, he adds, and large-population studies are now needed.

An investigation by the National Institutes of Health last year found no firm evidence for the existence of a 'Gulf War syndrome', and Bailer is noncommittal. He admits that there are "a significant number of veterans who are sick and suffering", but adds that the nature and specificity of their illnesses remained unknown.

Colin Macilwain

Ballot urged on ageing centre plans

San Francisco. Animal-rights and environmental activists have stalled the construction of a \$100-million centre for multidisciplinary research on ageing in California, claiming that plans for the centre are too ambitious.

The Buck Center for Research in Aging is the result of eight years of planning following a court decision that \$5 million of the funds of the Buck Trust should be devoted annually to an international centre on ageing.

The trust supports a variety of social services and environmental groups in Marin County, north of San Francisco. Bulldozers had begun preparing land for construction of the new center when opponents collected almost 19,000 signatures to put the project to a local ballot. Now county officials will have either to rescind their support of the Buck Center, or ask local citizens to vote on whether to build the project.

The centre hopes to attract scientists to carry out basic, clinical and social policy research on ageing. The first five years of work are planned to be devoted to the ageing brain, with projects in Alzheimer's disease, Parkinson's disease, cognitive problems and depression.

The board of the centre expects researchers to supplement its \$5-million annual budget with grants from the National Institutes of Health, support from private philanthropists and agreements with corporations.

But local environmentalists are objecting to the project's size and location. The centre owns 488 acres of land and the project would eventually include a 355,000-square-foot research and education complex, a commercial centre and about 120 condominiums.

Officials say that they plan to build on only 6 per cent of the land, and to landscape or otherwise alter slightly more than half the parcel. But local residents say that plans to stabilize the side of Mount Burdell (the mountain that the site backs on to) in order to manage radioactive and infectious waste safely would destroy an important greenbelt.

Animal-rights activists argue that the centre's plans to study only rodents or similar animals are likely to change. Others complain that the financial bonds needed to fund the centre could put other programmes funded by the Buck Trust in financial jeopardy.

Craig Perrin, coordinator for the Committee to Save Mount Burdell, which has led the protest campaign, says that the community would have preferred a centre focused solely on human clinical studies, regarding basic research as outmoded.

But a spokeswoman for the centre says that it has worked with the local community to find ways to mitigate the development's impact, and is confident that local citizens will approve the plan.

Sally Lehrman

Blood transfusion system comes under fire in Canada

Montreal. A group of international experts has criticized the operation of Canada's blood transfusion system and warned that safety could be "seriously compromised" unless deficiencies in its procedures are corrected. The experts were called in by a commission of inquiry — headed by Mr Justice Horace Krever — set up last spring to review the country's blood transfusion system.

In its report, the expert group claims in particular that the responsibilities of the organizations making up the system — the Canadian Red Cross Society, the Bureau of Biologics (the federal regulatory agency), the Canadian Blood Agency (the funding body) and various public health departments — are poorly defined, and cause conflict among the agencies.

The group, which was chaired by Kenneth D. McClatchley, professor of pathology at the University of Michigan, Ann Arbor, in the United States, also criticizes the overuse of blood and blood products, understaffing, and a lack of expertise at the Bureau of Biologics.

Its overall verdict is that Canada's blood system is as safe as that of other industrialized countries. But although procedures for collecting blood and manufacturing blood products at regional centres are "satisfactory", they are nonetheless "fragile".

The experts recommend a restructuring of the system to eliminate conflicts among the various agencies, and to define their responsibilities for ensuring safety. The Bureau of Biologics should be given more resources to enable it to monitor safety, adds the report, while the Red Cross needs to improve planning and management.

Furthermore, the scope of research should be enlarged, it says, from mainly investigator-led studies of basic medical problems to wider issues of health-systems research. In particular, research should include evaluating strategies to improve safety and efficiency in blood collection and delivery.

The Red Cross has accepted most of the group's conclusions, in particular those that would rebuild the public confidence in the system damaged, as in many other countries, by the contamination of blood supplies with human immunodeficiency virus in the mid-1980s.

The recommendation that regulatory procedures should be improved has also been supported by the Canadian Hemophilia Society. According to the society's president, Durhane Wong-Reiner, many senior employees at the Bureau of Biologics have quit because of frustration with its structure, although the bureau itself says it intends to double its staff to 24.

David Spurgeon

Computer networks win new funding in Japan

Tokyo. Despite the most austere budget in 40 years, Japan's Science and Technology Agency (STA) and the Ministry of International Trade and Industry (MITI) have managed to win large increases in funding for computer networks and computer-related facilities and projects in the budget for the fiscal year 1995, which was finalized at the end of last month.

The new money will result in a significant increase in Japan's contribution to the worldwide Internet system of computer networks, in which Japan has so far been slow to participate.

For the first time since 1955, Japan's overall budget will decrease — by 2.9 per cent — relative to last year because of shortfalls in tax revenue resulting from the prolonged recession. The STA and MITI have therefore done well to win even small rises in their budgets for research and development (see table). The budgets have to be formally approved by the Diet, the Japanese parliament, before coming into effect in April. But in practice they are unlikely to undergo any change.

The STA has obtained an increase of nearly 50 per cent in funding for its interministry computer network, already under construction, which will have a high-capacity backbone network linking Tsukuba science city, Tokyo and the Kansai region encompassing Kyoto and Osaka. The network will be linked to other computer networks in universities and national research institutes run by various ministries, and will also have a high-capacity overseas link.

STA has also obtained ¥2,776 million (US\$28 million) for a high performance computer system to be used by its various research institutes through the computer network to model phenomena such as molecular interactions and global climate change. The budget includes funds for both software and a supercomputer, although STA has yet to decide where the supercomputer will be located.

MITI has been granted about ¥1 billion (US\$10 million) for computer databases and a new research information centre in Tsukuba science city that will provide outside users access to the ministry's powerful supercomputer centre in Tsukuba.

Until now, the supercomputer centre could only be used by the ministry's 2,500 researchers. But the new centre will be open to users in universities and industry who are participating in MITI projects.

One such project is the Real World Computing (RWC) project, a follow-up to the fifth generation computer project, which gets a 20 per cent boost in its budget, to ¥6 billion. The government's goal is to develop massively parallel computer systems using various approaches, such as optical computing and neural networks.

There is also a significant increase in the budget for STA's human genome project, which will grow to ¥2.4 billion. Japan's genome researchers were among the first to establish computer networks in Japan in the late 1980s, and will no doubt benefit from the rapidly expanding networks.

David Swinbanks

HIGHLIGHTS OF JAPAN'S R & D BUDGET

(in billion yen: US\$1 = ¥100)

	1995	% change
Ministry of International Trade and Industry		
Total R & D budget request	295.4	+3.8
Industrial Scientific Technology	24.9	+5.4
New Sunshine Project	54.2	+2.7
Real World Computing Project	6.0	+20.5
Intelligent Manufacturing System Project	1.3	+0.1
Global Environment Research	12.5	+4.2
Unmanned Space Platform	8.3	-0.3
Research Databases	0.8	+46.9
Research Information Center (new)	0.8	-
NEDO Fellows (new)	0.4	-
NEDO International Grants	0.9	+3.4
Japan Key Technology Center	26.0	0.0
Science and Technology Agency		
Total R & D budget	646.1	+6.8
Special Promotion Funds	18.5	+19.3
Interministry computer network (1.6)		+45.5
High Performance Computing	2.8	+1508.7
Space	163.9	+5.0
Nuclear Power	342.3	+5.7
(ITER)	8.3	+3.7
(SPring-8)	15.1	+37.3
Human Genome	2.4	+26.3
Human Frontier Science Programme	3.5*	-2.8
Frontier Research System	3.9	+19.9
STA Fellowships	2.2	+15.8

* includes ¥1.4 billion from MITI

UK researchers highlight lack of funds for overheads

London. David Hunt, Britain's minister for science, last week promised to raise the question of whether English universities are using enough of their funds to cover the overhead costs of biomedical departments with Gillian Shephard, the secretary of state for education, and with the Higher Education Funding Council for England (HEFCE).

The promise was made in response to concern voiced by 12 biomedical research scientists, all funded by the Wellcome Trust, who have come (or returned) from abroad to work in Britain. They attended a meeting with both Hunt and Virginia Bottomley, the secretary of state for health.

The main purpose of the meeting was to discuss why, at least in some research fields — and with the support of private funding — Britain can demonstrate a clear 'brain-gain' that stands in contrast to the highly publicized 'brain-drain' of recent years.

But both the scientists and the officials of the Wellcome Trust, which this year expects to provide around £200 million for research, also used the occasion to tell Hunt

of the problems that arise from the fact that universities receive money from the government through the funding councils on a block grant basis.

The amount of money awarded to a particular university depends on many factors, including its research income and the quality of its research. But a university is free to spend the money it receives as it chooses.

Julian Jack, a governor of the Wellcome Trust, says that some universities do not pass enough of this money on to the departments that attract the research income. As a result, where charities such as the Wellcome Trust — which do not provide overhead costs — fund the research, a lack of infrastructure support provided by the university means that departments are running at a deficit, or curtailing their research.

Professor John Bell, a Canadian who moved to Britain from Stanford University in California, and who is now chairman of the Nuffield department of clinical medicine at the University of Oxford, says that the lack of adequate government support for

research overheads is "singularly the biggest problem" facing most university biomedical research departments.

Nevertheless, the scientists who met Hunt and Bottomley listed several attractions of biomedical research in the United Kingdom. These included the availability of long-term support — particularly in grants from the Medical Research Council (MRC) and the Wellcome Trust — and the opportunities for collaboration resulting from the proximity of Europe.

Other factors include a smaller and more cohesive research community than that in the United States, and the fact that the National Health Service makes it easy to combine academic research with clinical work and access to patients and patient records.

But, despite the advantages, all the scientists who spoke to Hunt agreed that a major disadvantage is the lack of international competitiveness of UK salaries. According to Bell, this is symptomatic of the low status attached to scientists in the United Kingdom.

Maggie Verrall

Roche pledges independence to new institute

Washington. The Swiss pharmaceutical company Hoffman-La Roche, still facing controversy over its decision last November to transfer the renowned Roche Institute of Molecular Biology (RIMB) from Nutley, New Jersey to Stanford, California, has promised that the relocated institute will continue to enjoy freedom to pursue basic research.

Roche is still trying to find a director for the new institute — which will also have new staff and a new mission in genomics — but is “not going to interfere after that”, says Jurgen Drews, the company’s president of international research and development. He promises to keep the charter that guaranteed the independence of the old institute.

But scientists at Nutley — almost all of whom are likely to leave, despite Roche’s suggestion that they can “discuss the possibility of remaining with the RIMB” when the new director is appointed — say that the concept of the institute’s independence has been permanently shattered by Roche’s actions. “Our charter was supposed to protect us from short-term changes in corporate outlook,” says one senior researcher who asked not to be named. “It didn’t.”

The institute was established by Roche at Nutley in 1967, and has developed a worldwide reputation through its successful cloning of interferon and a host of other scientific advances.

Many of its researchers claim that the decision to move the institute is the result of a combination of Drews’ personal preference for California and its sunshine over the New Jersey suburbs of New York, as well as a reshuffling of research resources following Roche’s \$5.3-billion acquisition of the Stanford-based drugs company Syntex.

The closure of the Nutley institute will be counted, the researchers say, among 5,000 job cuts that Roche said would come from post-merger rationalization. Severance costs at Nutley will be itemized in Roche’s accounts as Syntex acquisition costs.

But, given that changes were in the wind even before the Syntex deal, another powerful element in Roche’s decision appears to have been its failure, after more than a year of trying, to attract a suitably qualified geneticist to replace Herbert Weissbach, who is retiring as director of the institute. Drews has wanted to reorientate the institute towards genomics for some time, and is now hoping that the promise of hand-picked staff at a new site will make it easier to attract the right person.

Researchers at Nutley are angry that what they regarded as their tenure at the institute has been abrogated, even though they have been offered a severance package of one month’s salary for each year of service, as well as transition grants to enable their research to continue until they find new sources of support.

But Shirley Tilghman, a professor of life science at Princeton University who chairs an external board of scientific advisers to the director of the institute, describes Roche’s offer as “fair and reasonable”, and says the board has tried to ensure the well-being of RIMB scientists. She admits that the sudden decision to move the institute “was shocking to those of us who live and work in academic science”. But Tilghman feels that Roche was acting “within its rights” in deciding to change the focus of the institute.

The new focus will depend largely on the new director. “We’re looking for someone who is interested in developing a vision of what such a new institute could be,” says Drews. But he promises that the Stanford centre will not just be another genomics centre. “It is post-genomics: the science

after we have the sequence of the genome,” he says, adding that funding of the institute will remain about \$20 million a year.

Drews says that the concept of a broadly based molecular biology institute made sense in the 1960s, when the subject was relatively new. “Today it is different, it is mainstream science. The mission is a little out of date, and we felt it needed a change.”

But scientists soon to lose their posts at Nutley feel insulted by Drews’ logic, as they see themselves already at the forefront of efforts to understand the function of genes. “They [Roche] don’t seem to understand what genomics is,” says one. “You need genomics people to sequence the genome, and to do gene function work you need broadly based molecular biologists.”

Colin Macilwain

Congress seeks regulation controls

Washington. The new Republican-led US Congress moved swiftly last week on its pledge to reduce the size and influence of the federal government, introducing three bills on the session’s opening day designed to limit ‘unfunded mandates’ imposed on states, local governments and businesses.

These ‘mandates’ refer to environmental, safety, health and other regulations that are required by federal law but do not include federally appropriated money for localities to carry them out.

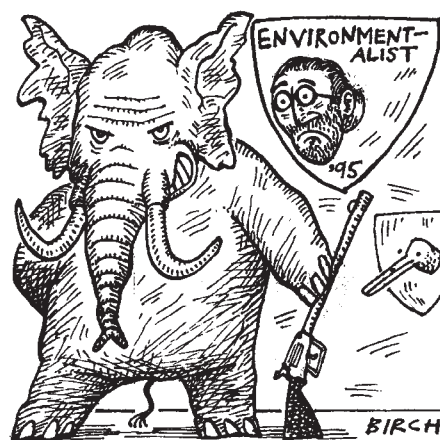
Republicans have declared the elimination of unfunded mandates to be one of their priorities, and a bill drafted by Senator Dirk Kempthorne of Idaho was the first introduced in the Senate. But the issue is not strictly partisan. The proposed legislation has Democratic co-sponsors, as well as the support of President Bill Clinton.

Furthermore, although environmentalists fear that this marks the beginning of a larger campaign to weaken environmental legislation in the new Congress, the proposed bills are less radical than demands in the Republicans’ *Contract with America* and elsewhere. For example, they would not apply to laws already on the books — a particular fear of environmentalists.

According to the proposed legislation, if Congress imposes a mandate, such as a new drinking water standard, on a state or local government, it must also appropriate funds. If federal money is not provided, the mandate must be scaled back to match available funds, or Congress must take a separate vote to impose the mandate.

Any new law expected to place a significant financial burden on local governments or businesses would require the Congressional Budget Office to estimate those costs before the law is passed.

Estimating the cost of new regulations is



notoriously difficult, and making such estimates mandatory could therefore bog down the regulatory process. Restricting unfunded mandates may also lead to fewer federal regulations, as Congress would have to pay more of the regulations’ costs.

If federal mandates are scaled back, some state and local governments may relax their own environmental regulations to attract businesses, becoming what Blakeman Early, a Washington consultant and former lobbyist for the Sierra Club on clean-air legislation, calls “pollution havens”.

Environmentalists will face bigger challenges in this new Congress, but the issue of unfunded mandates carries much symbolic weight with state governors and mayors, who have long complained about them. Republicans are eager to win the governors’ support for an amendment to the US constitution requiring a balanced federal budget; such an amendment would have to be approved by three-quarters of the states. The bills introduced last week are therefore expected to pass quickly, perhaps by the end of the month.

Tony Reichhardt

Bavaria seeks opt-out from university funding system

Munich. The *Land* (state) of Bavaria has said that it wants to opt out of Germany's communal funding system for supporting university infrastructure because the federal government in Bonn does not want to contribute more to the fund than it did last year.

In principle, the Bonn government's decision prevents Bavaria from implementing its full university building programme, even though, as a relatively rich state, it can afford to do so.

Bavaria's declaration has raised a question mark over the future of a 25-year-old law that splits the costs of financing the construction, maintenance and expansion of university buildings — as well as the purchase of major items of university research equipment — equally between the federal government and the host *Land*.

In general, higher education is the responsibility of Germany's 16 *Länder*. But the 1970 law was designed to promote common standards by ensuring that buildings in poorer *Länder* would not become run-down through a lack of money. Building and equipping proposals put forward by the *Länder* are considered by the Wissenschaftsrat, Germany's science council.

Normally the council's recommendations on financing levels are accepted by all sides. This year, however, general budgetary pressures meant that the federal government agreed to pay only DM1.8 million (US\$1.15 million) for 1995, even though the Wissenschaftsrat had said that Bonn should contribute at least DM2.3 million (see *Nature* 371, 371; 1994).

The Wissenschaftsrat expressed concern that this capping of funds would prevent

universities from making any new plans in the next few years. Bavaria, which by law should only match its share of the federal funds, has now indicated that it does not wish to be constrained by the financial problems of the federal government, nor concerned by those of other *Länder*.

Bavaria is one of Germany's richest *Länder* and is keen to become the country's science capital (see *Nature* 371, 189; 1994). Faced with rapidly increasing student numbers, it has embarked on a major university building programme to help relieve overcrowded classrooms.

The *Land* has already been successful in attracting private finances for some university building projects, including a centre for applied solid-state physics in Augsburg, and a new building for the faculty of applied science in Bayreuth. But Bavaria's prime minister, Edmund Stoiber, wants the government either to increase public funding for universities or to release the *Land* from the current joint-funding agreement.

Last month, Stoiber announced that Bavaria would continue with its full building programme, as approved by the Wissenschaftsrat. He said that the *Land* will pay the full costs, and then fight retrospectively in the courts for the federal share, to which he claims it is entitled.

If it loses the battle in the courts, said Stoiber, Bavaria will start negotiations to pull out of the agreement altogether, and fund all university building itself. But in such circumstances, it would expect the federal government to direct a higher share of taxes to Bavaria in other forms.

Alison Abbott

Russian academy backs off take-over of applied research labs

Moscow. Leaders of the Russian Academy of Sciences (RAS) have withdrawn their support for plans to create a new organization, the Russian Science and Technology Academy (RSTA), which would have taken over responsibility for all publicly funded applied science.

The plans were announced last autumn (see *Nature* 372; 208; 1994), and would have given the RSTA control over research institutes, factories and scientific organizations that are now under the jurisdiction of the State Committee of Defence Industries, the Ministry of Atomic Energy and other ministries involved in research and development.

One goal of the proposed move was to save Russia's military research and development complex from collapse. But at a recent meeting of the academy's presidium, it became clear that the 'superacademy' had more opponents than supporters.

The reasons behind the presidium's climbdown on the issue remain unclear. One factor may have been the resignation of Oleg Soskovyets, the first Vice-Premier of Russia, who issued the official order setting up the academy last year.

But the meeting was also told by Igor Lebedev that scientists in Siberia were strongly opposed to the creation of the RSTA, which many regard as potentially increasing bureaucratic controls, while Gorimir Chyorny questioned whether an academy is the best type of organization to manage applied research institutes. Another academician, Alexander Yanshin, told the meeting that he opposed the creation of any form of state-run academy.

Boris Saltykov, the Russian science minister, has already expressed his opposition to the creation of the RSTA, especially if this were to involve the transfer of control of public institutes. Saltykov's views were endorsed at the academy meeting by Andrei Gonchar, the RAS vice-president.

Indeed, several applied research institutes have themselves opposed the plans. For example, Fyodor Reshetnikov, the director of the Institute of Organic Chemistry (part of the Ministry of Atomic Energy), has questioned assurances by another academy vice-president, Evgeny Velikhov, that institutes would not be forced to join the RSTA. He claims that his institute was included in the plans without his consent.

The RAS's decision to withdraw its support for the RSTA does not necessarily mean that the project will be abandoned. But the continued opposition of the minister of science, combined with the RAS's about-turn, makes its realization increasingly unlikely.

Vladimir Pokrovsky

Cambridge turns to its gospels

London. The sixth-century **Canterbury Gospels** (see right) is one of the highlights of an eclectic mixture of books, photos, instruments and memorabilia that went on show at Christie's auction rooms in London last week. The exhibits are owned by the University of Cambridge and are intended to demonstrate the intellectual contribution made by the university since it was founded more than 700 years ago.

The exhibition is part of the Campaign for Cambridge, a 10-year fund-raising drive, launched five years ago. But none of the exhibits is for sale and admission is free; university officials hope that publicity generated by the display, opened last week by Princess Margaret, will give a new boost to fund-raising efforts.

M.V.



Belfast university promises commitment to neutrality . . .

Paris. Northern Ireland's oldest and largest university, the Queen's University of Belfast, is to discontinue the playing of the United Kingdom's national anthem, 'God Save the Queen', at graduation ceremonies, and to replace it with Beethoven's 'Ode to Joy', the anthem of the European Union.

The change is one of a number of measures being taken by the university to promote a "neutral working and social environment". These mark a major shift in the university's position on the conflict in Northern Ireland, according to Professor Lesley Clarkson, who is senior pro-vice chancellor, and also chairman of the university's equal opportunities group.

The measures were adopted by the university's Senate last month on the recommendation of a special advisory group set up as a result of an independent investigation in 1993 into employment practices at the university, which confirmed that the university had discriminated against Roman Catholics and also prompted a review of recruitment practices (see *Nature* 363, 475; 1993).

As a further symbol of its desire to demonstrate its neutrality, the university Senate also decided that the band of the Royal Ulster Constabulary — the largely Protestant police force which has traditionally provided entertainment at graduation garden parties — "should be seen as one of a number of suitable bands, rather than the only one".

The significance of these changes stems from the fact that the terms 'Protestant' and 'Catholic' denote two communities with different political aspirations; much of the Catholic minority favours a United Ireland, while much of the Protestant majority wants Northern Ireland to remain part of the United Kingdom.

In recent years, the number of Catholic students at Queen's has overtaken that of Protestants, and official signs in the student's union are now displayed in both

English and Irish. The special advisory group points out that some students perceive this practice as 'divisive', and recommends that such signs be "used and identified as cultural rather than political symbols."

The advisory group also recommends that the university should promote both the 'Ulster/British' and 'Irish' cultures, and also take an active role in promoting peace. Encouraging cosmopolitanism — and a European dimension, in particular — in all university activities, it adds, may help to diffuse tensions by broadening "perspectives" beyond the "narrow confines" of Northern Ireland.

Such changes represent an important shift for a university that Clarkson admits has "held itself aloof from the troubles" for a quarter of a century. Indeed, those working on the campus, which is sited in the leafy suburbs of south Belfast, have been able to remain oblivious to, or at least sheltered from, the widespread violence in, for example, the terraced-house ghettos less than 2 km to the west.

But Clarkson also admits that "hanging on to divisive emblems is not furthering the work of the university". He says that the university's plans to reassert itself politically have been given added momentum by the current peace negotiations, which were prompted by the announcement of a ceasefire last year by both the Catholic Irish Republican Army (IRA), and Protestant terrorist groups.

The university's new policies have attracted the wrath of Protestant politicians. James Molyneux, the leader of the moderate Ulster Unionist party, for example, last week rejected an honorary doctorate from the university in protest at the university's decision to scrap the national anthem. Clarkson says he is "saddened" but "not surprised" by the attitudes of Protestant parties.

Declan Butler

Patent changes in India clear path for US accord

New Delhi. India last week issued a presidential decree proposing that the country's patent law be amended to allow patents on novel products. At present, legislation passed in 1970 authorizes patents only on processes, allowing India to avoid paying royalties on drugs and chemicals patented elsewhere simply by developing new processes for their production.

The presidential decree, according to a US official, will clear the way for the signing of an 'omnibus' science agreement with the United States, first submitted to the Indian government for approval two years ago, but left to gather dust because of the disagreement over patent legislation.

India's current patent laws have enabled it, for example, to manufacture the chlorofluorocarbon (CFC) substitute without paying royalties to the company Dupont, which holds a patent. But the country has been obliged to make the change as part of its acceptance last year of the General Agreement on Tariffs and Trade.

The change will be welcomed by the United States, which estimates that poor patent protection in India has caused enormous damage to US trade.

But not everyone is happy about the new decree. Critics argue that any gains in health and food production will be offset by large increases in the price of drugs and agrochemicals. Furthermore, Indian drug companies stand to lose much of their 60 per cent share of the domestic market to multinational companies.

Observers outside the drug industry agree that the lack of patent protection on products has stifled innovation. Indeed, the number of patents granted in India fell from more than 40,000 to around 12,000 between 1968 and 1988. Indian innovators have also tended to patent their inventions abroad.

The change will boost research and development, says Mahesh Prasad Bhatnagar, a legal expert at the National Research Development Corporation, the body responsible for exploiting domestic technologies.

Although the new patent regime will pit small Indian companies against multinationals, it will also contain safeguards to protect public interest, such as giving greater discretionary powers to the patent office. India will also benefit from a 10-year transition period agreed with the World Trade Organization.

But before the decree becomes law it must first be ratified by both houses of parliament. The opposition parties are divided on the issue. But the ruling Congress party is in a minority in one of the houses, and observers predict that the amendment will have a rough ride.

K.S. Jayaraman

. . . as US aims to boost science links

Paris. The United States is expected to announce a package of measures to promote the transfer of US research and technology to Northern Ireland at a special conference for trade and investment in Ireland which is to be held in Philadelphia in April.

The package is being developed by a working group set up jointly by the US Department of Commerce's Technology Administration and the Industrial Research and Technology Unit of Northern Ireland's Department of Economic Development. It is expected, in particular, to

provide the province with favoured access to US technology and to link Irish science centres with corresponding organizations in the United States.

The US administration has said it considers such cooperation vital to "supporting peace" in the province by encouraging economic growth. In a further sign of confidence in the peace process, the company Du Pont announced last month that it is to invest £13.5 million (US\$ 21million) on developing prototype textile technologies at its Lycra facility in Derry.

D.B.

Ethics board members quit over 'imbalance'

Munich. The biologist and Nobel prizewinner Rita Levi-Montalcini last week resigned as a member (and honorary president) of Italy's national bioethics committee, in protest at a decision by the former prime minister Silvio Berlusconi to appoint nine new committee members, all of whom are orthodox Roman Catholics.

Two other members, including the committee's vice president, have also resigned. The nine appointments create a strong Catholic majority to the committee — and have been widely condemned as an act of political manipulation.

The national bioethics committee was established in 1990 as an advisory group to the government, in line with similar committees in other European countries. Its 40 members are appointed by the prime minister, and its composition is intended to represent a balance of the different views on bioethics that prevail in Italy.

Berlusconi, the leader of the Forza Italia party, upset this balance when he signed a decree on 16 December, during the last days of his government, nominating the new Catholic members. The decree also gives the committee a four-year term of office, in contrast to its previous two-year terms.

Berlusconi hopes to be asked to form a new coalition government and his creation of a majority of orthodox Catholics on the

bioethics committee is widely seen as an attempt to woo potential new coalition partners by appealing to the Roman Catholic Church and thus the Partito Popolare Italiano (PPI) party. The PPI is the successor to the disgraced Christian Democrats. It is not at present represented in government, but could be if new elections are called.

The move is also being seen as an attempt to appease Forza Italia's main coalition partners, the right-wing Alleanza Nazionale. This wants to change Italy's abortion laws, which permit abortion in certain circumstances.

Such political manipulation makes nonsense of the committee's attempts to represent a balance of views, says its former vice-president, Giovanni Berlinguer, professor of public health at the First University of Rome's faculty of science. Berlinguer announced his resignation last week, at the same time as Levi-Montalcini and a second committee member, Eugenio Lecaldano. "Some of the most prominent scientists and philosophers were removed from the committee to make way for new members from the very dogmatic, orthodox Catholic faction," he says.

One scientist who has been dropped from the committee is molecular biologist Glaucio Tocchini-Valentini, who first learnt about his dismissal in the newspapers several days

before he received his official letter from Berlusconi.

The resignation of Levi-Montalcini has provoked particular concern among Italian scientists. Berlinguer hopes that the resulting publicity will force the next prime minister to return the committee to "a more sensible balance", and says that in such circumstances he would be prepared to withdraw his resignation.

But Levi-Montalcini is not likely to return. She is keen that her resignation should not be seen as anti-Catholic. "I resigned because the balance of the committee had been changed; I would also have resigned if the balance had been changed in the other direction," she says.

But she is also unhappy with the work of the committee, saying that it has moved too slowly and keeps repeating themes — such as the ethics of *in vitro* fertilization — that have already been dealt with and legislated for in the United Kingdom and France, without drawing conclusions.

Berlinguer agrees that the work had been slow, but says that the dialogue has always been good. He adds that "in order to influence, it is necessary that the legislative powers are ready to be influenced. But bioethics has never been of great concern to the Italian government."

Allison Abbott

'Out-dated figures misled critics of space station'

Washington. The National Aeronautics and Space Administration (NASA) has hit back at the report of a government watchdog agency which said that it will lack sufficient scientists on the ground to support the research it is planning to carry out on the space station.

The space agency says it has already started to support most of the microgravity and life sciences investigators that the General Accounting Office (GAO) has claimed it will not have in place when space station research starts in 1998.

The GAO report, *Space Station: Plans to Expand Research Community Do Not Match Available Resources*, says that NASA will need to increase the number of ground-based microgravity researchers it supports from 73 in 1992 to 240 by 1998, and concludes that this cannot be done within planned funding levels.

Opponents of the space station led by Senator Bill Cohen (Republican, Maine), who requested the report, jumped on its central conclusion when it was released late last month. "With its research budget frozen, NASA will not be able to maintain its current number of scientists, much less triple their number as the space station re-

quires," Cohen said.

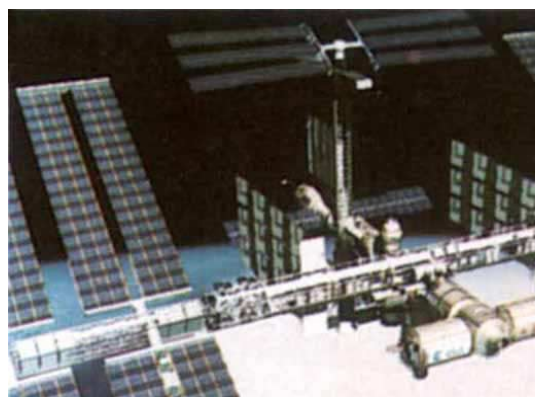
But Arnauld Nicogossian, a deputy administrator of NASA's office of life and microgravity science, says his office had already built up strength since 1992 to the point where it is funding 209 ground-based microgravity investigators. "The GAO used the old data," he says. "I don't know why".

Frank Degnan of the GAO, the investigative arm of Congress, says that NASA officials — including Nicogossian — reviewed the draft report at a meeting in October, and should have raised any problems then. "No-one indicated that the number had grown to 209," he said. "If they weren't aware of it then, they should have been."

But NASA officials say the GAO threw so much information at them during the October meeting that they overlooked some figures that were correct, but incomplete. Only later did the figures become a focal point of the report.

Bob Rhome, director of NASA's microgravity science and applications divi-

sion, says that while the data in the GAO report are technically correct, their interpretation was wrong. Rhome says that NASA has been able to build up microgravity



The space station will need considerable back-up from ground-based research teams.

research quickly over the past two years by diverting resources from the space station's research and development budget — and has informed Congress of what it is doing. "The fact is that we've got four more years to find 31 extra investigators," says Rhome.

Colin MacIlwain

More on DNA typing dispute

SIR — Lander and Budowle¹ imply that the inferential problems of forensic genetics have been solved and that further studies are not warranted. Perhaps a spectator, with an interest in the more difficult, but equally confused, problems of linkage analysis could comment. The arguments of the authors and their associates instructed and entertained readers of *Science* two years ago in a series of articles when the differences in the misunderstandings of the various contestants were to some extent clarified. They have recently been catalogued and dissected in detail by Morton². Different misunderstandings are hardly an adequate justification for gratuitous advice.

Forensic genetics covers a wide field, sometimes involving degraded specimens, or inferences on genetic mixtures on good samples (paternity tests), or the resolution of mixtures in degraded mixtures (rape). If there is an adequate supply of fresh blood, the various shot-gun methods pioneered by Jeffreys provide an unambiguous answer, but, as with ordinary fingerprints, convey too much information for mathematical analysis and are simple enough for judges and jurors to comprehend and interpret without advice. The quantity and quality of the blood available in the O. J. Simpson case was not stated.

If samples are limited or degraded, techniques involving amplification of short segments of DNA are necessary, and their analysis involves the summation of the evidence from each part. While it is obvious if any result differs, difficulties arise when they do not. This is the central problem of classification, and indeed of language, and can hardly be dismissed as a non-problem in a few pages.

Even the simplest model presents formidable difficulties. Suppose we have a bag of coins, one of which is double-headed. A coin is removed and tossed, and a decision with consequences of life or death has to be made on the result of a defined number of tosses of a single coin selected at random. If any tail appears, the problem is solved. If not, the odds against the coin selected being a regular coin after ten tosses are about a thousand to one. However, if there are a thousand coins in the bag, the odds against this coin having been selected are also a thousand to 1. It is not possible to give judgement without knowing the size of the bag, as Laplace observed in a similar context. In this case the regular coins are unbiased. In the forensic case of murder they are not. Murders preferentially involve relatives and neighbours.

Lander and Budowle's argument appears to be that if the bag were big enough to carry, and the coins of a regular variety, then we could assume a maximum

size, or ceiling, and just keep tossing for as long as necessary. This is obviously true, but the 'ceiling' is so arbitrary that it can hardly 'support' any very elaborate inference. Even this has problems, for the 'regular' coins are not unbiased, as the population of the world does not consist of unrelated individuals: we are all relatives and most of us have several close relatives.

This introduces even graver problems if segments of chromosomes, as well as murderers, are assumed selected at random. A solution that could imply, at odds ratios comparable to the population of the world, that no two persons would have any realistic chance of being identical seems seriously flawed.

The courts will have enough problems in the O. J. Simpson case. It would be unfortunate if public appreciation of population genetics, a subject largely developed in the first half of this century, were to become one of comic denigration and add support for creationism and genetical brands of political correctness.

J. H. Edwards

University of Cambridge,
Genetics Laboratory,
Department of Biochemistry,
South Parks Road,
Oxford OX1 3QU, UK

SIR — Although the National Institute of Justice (NIJ) does not, as a general rule, take positions on issues of this sort, we wish to clarify one misunderstanding in the article by Lander and Budowle¹. They write that: "The NRC [National Research Council] — at the urging of the National Institute of Justice, representing the academic wing of forensic scientists — has concluded that the best solution is to constitute another *ad hoc* committee on DNA fingerprinting, composed primarily of statisticians and population geneticists".

First, it is not the National Institute of Justice that urged the NRC to convene this board; the NRC jealously guards its independence. Several people suggested that a new committee be convened, including the director of the Federal Bureau of Investigation (FBI), William Sessions, in a letter to the NRC. NIJ was not one of those, but when the NRC decided to convene this committee, it approached NIJ for funding. At the urging of the FBI, and others, NIJ agreed to provide much of the funding for the new committee.

Second, the NIJ does not represent "the academic wing of forensic scientists". It doesn't represent any group of forensic scientists, academic or otherwise. It supports well-designed research into the forensic sciences by practitioners in crime laboratories, academics in universities and others, including work in the labor-

atories of federal law-enforcement agencies. Every peer-review panel at NIJ includes practitioners who ensure that NIJ research meets the real needs of US crime laboratories at the local, state and federal levels, as well as uniquely qualified experts from academic life and the federal and military forensic laboratories.

The role of the National Institute of Justice is unique. It serves as an independent research agency supporting all levels of the law enforcement and criminal justice system, from the local to the federal. It has, for nearly a quarter of a century, been the principal source of federal funds for the forensic sciences community and takes very seriously the legislative directive that it serve the practical needs of the law enforcement and criminal justice communities.

David G. Boyd

(Director, Science and Technology)
National Institute of Justice,
Washington, DC 20531, USA

SIR — Having been involved in several trials using DNA typing, I wish to reply to Lewontin and Hartl^{3,4}. Lewontin states "...juries are no more capable of understanding probability statements than they are of interpreting any other piece of highly technical information. . ."³.

Juries have been coping with probability statements with respect to serological typing with judiciousness and effectiveness for decades. DNA testing is not qualitatively different from serology. Perhaps Lewontin is reacting to the phenomenon that, with rare exceptions, judges and juries who have listened to his railing against DNA testing have chosen to be persuaded by the opposition's point of view. This does not prove that juries are incapable of understanding the issues, but merely attests to the lack of persuasiveness of Lewontin's arguments³.

It is impossible to determine with certainty the genetic group or subgroup of any accused individual. Even in the rare instances when extensive pedigree information is available, experience with paternity testing has demonstrated that a significant fraction of paternity is misassigned. The use of the ceiling principle ensures that the suspect will be afforded the maximum conservatism with respect to the probability estimates and should not be considered an 'interim' solution.

I agree with Lewontin that the refusal by the FBI laboratory of outside inspection and data verification is troubling, especially when I have been called upon to testify in support of its findings. Regardless of the reasons for this policy, I believe that the FBI laboratory should be held to the same standards and requirements as other laboratories.

The term DNA fingerprinting, as I understand it, refers to a patented process of Cellmark Diagnostics involving multi-

locus probe testing. Therefore, as Lewontin points out, this term is used incorrectly throughout the Lander-Budowle article.

The continued existence of a Flat Earth Society and the increasing popularity of Creationism demonstrate that it is never possible to convince every individual of the validity of a scientific theory. However it is clear that the concepts of evolution and the spherical shape of our planet are "generally accepted" in the scientific community and would pass the Frye test for courtroom admissibility. *Nature's* chronicle of the arguments against HIV as the causative agent of AIDS is another example of how a tiny, vocal minority with access to media outlets can attempt to sway public opinion against generally accepted medical and scientific opinions.

Charles M. Strom

Section of Medical Genetics,
Illinois Masonic Medical Center,
836 Wellington Avenue,
Chicago, Illinois 60657, USA

SIR — Lander and Budowle¹ highlight legitimate domains of convergence between two former opponents but stint on views of others and on unresolved issues. Lander and Budowle strongly doubt that the new NRC committee can make recommendations substantially to improve forensic DNA analysis. Recent correspondence by Lewontin³ and Hartl⁴ is largely tangential to that issue. Their letters consist mainly of speculation on the motives of individuals and the possible future behaviour of the FBI. Helpful comments by Lewontin on possible improvements in quality control and blind testing are diluted by other comments, for example his patronizing assertion that jurors are incapable of understanding the meaning of a 1 in 4 probability and his insistence that the situation is basically hopeless until an entirely different system for DNA identification is developed. New technologies for DNA identification are being developed and each will probably share some of the same problems in the current technology. Therefore, we need not wait for the millennium to find practical improvements. "A steady succession of *ad hoc* committees"¹ is undesirable, but significant work remains for the new NRC committee in advancing the way the existing technology is applied.

(1) Exceedingly small genotype frequencies (for example $<10^{-6}$) may be calculated and, to make the number smaller, one simply has to type more polymorphic loci. Such probabilities are presented to jurors who assess their meaning as best they can and with the assistance of experts such as Lewontin, Hartl and ourselves. However, it is fruitless, beyond a certain point, to continue to type additional markers when we are already as certain as we can be, based on *one* valid test, of genotypic identity. Lander and Budowle cite a

frequency reported in one case, 1 in 738×10^{12} , as unrealistic, but provide no mechanism whereby the introduction of such a probability in a courtroom setting would be prevented or made sense of. Due to the possibility of error, exceedingly small genotype frequencies (say 10^{-7}) tell us little more than rare genotype frequencies (10^{-5}), but they may have prejudicial impact. It is more accurate to estimate a meaningful level of significance ($P < 10^{-4}$ or $P < 10^{-5}$).

(2) The first NRC committee suggested that genotype frequencies should be introduced with an error rate. Most practitioners of the forensic DNA art readily admit the possibility of error. Unfortunately, error rates are usually unavailable. Our suggestion in (1) would also mitigate this problem.

(3) Intrinsic to DNA testing are unique possibilities for eliminating error or fraud. We have two suggestions: (a) Different internal standards should be added to each sample to reveal sample mixing or mixups. (b) The individual performing an analysis should be unaware of which sample, out of a small group, derived from the suspect. This conforms to the established principle of blind testing.

(4) When, as frequently happens, multiple suspects are tested, the estimated match probability must be adjusted to take into account multiple testing. The NRC committee should also develop guidelines for the use of large databases of DNAs from criminal suspects.

(5) Special circumstances warrant the abandonment of the genotype frequency as the match probability. If individuals with a high degree of kinship have not been ruled out as the perpetrator, then the probability of the match is not the genotype frequency (and *pari passu*, idiotyping by DNA sequencing, as suggested by Lewontin, might exacerbate this problem). People differ in the number of close relatives they have; some have many close relatives, and inbreeding can enhance genetic identity by descent. Many individuals have half- or full-siblings unknown to them. As frequencies become increasingly remote, remote considerations loom increasingly large.

(6) What is the relevant genotype frequency, that of the evidence or that of the suspect?

(7) The ceiling principle method was formulated to account for possible differences in allele frequencies between populations. The second NRC committee should emphasize that the same consid-

erations universally apply, for example to DQ α .

The first NRC committee provided sound and conservative methods. Although conservative, the ceiling principle is arbitrary. Therefore it is doubtful if it would ever have been implemented save with the imprimatur of a distinguished committee. So far only the modified ceiling principle has been used because the systematic sampling of populations suggested by the NRC has not been performed. The second NRC committee is now in a unique position to refine the use of forensic DNA testing in important ways and to reexplore useful suggestions made by the first NRC committee but only partially implemented.

David Goldman

Jeffrey Long

National Institute on Alcohol
Abuse and Alcoholism,
Intramural Research Program,
Rockville, Maryland 20852, USA

SIR — Lewontin and Hartl^{3,4} complain that "because juries are no more capable of interpreting probability statements than they are of interpreting any other piece of highly technical information, there are insuperable barriers to their use in the courts". Perhaps I should recall the words of E. M. East, the pioneer quantitative geneticist⁵ commenting on Edgar Allan Poe: "as a poet and mathematician, he would reason well, as a mere mathematician he would not have reasoned at all." I am not surprised that lay people may be confused when some the terms used by Lewontin are also ill defined. The word 'idioplasm' was coined by Karl Wilhelm Nägeli⁶ before the Mendelian concepts became known and he used it in the sense of the entirety of the hereditary material. The newly developing genetics, after the turn of the century, abandoned this term for the more meaningful gene and genotype. Immunogeneticists revived it in the form of idiotope, the antigenic determinants in the variable chains of the immunoglobulins and idiotype as a collection of idiotopes distinguishing one type of antibody-producing cells from other clones of cells. Thus it is not a concept of DNA but of a protein and this is worth remembering even now, 30 years after synonymous codons became known. Thus, obviously it is not correct to call fingerprints — and I do not mean DNA fingerprints — idiotype(s). Also, it is well documented that some kindreds display no dermatoglyphs⁷. In some instances, forensic genetics cannot rely with absolute certainty on dermatoglyphics because of developmental differences, mosaicism and more than single gene involvement in the pattern⁸.

G. P. Rédel

3005 Woodbine Court,
Columbia, Missouri 65203, USA

1. Lander, E. S. & Budowle, B. *Nature* **317**, 735–738 (1994).
2. Morton, N.E. *Eur. J. Med. Gen.* **1**, 172–178 (1993).
3. Lewontin, R. C. *Nature* **372**, 398 (1994).
4. Hartl, D. L. *Nature* **372**, 398–399 (1994).
5. East, E. M. *Bot. Gaz.* **57**, 239 (1914).
6. Nägeli, K. W. *Mechanisch-Physiol. Theorie der Abstammungslehre* (München, Oldenburg, 1984).
7. Baird, H. W. *Lancet* **II**, 1250 (1968).
8. Slatis, H. M. *et al.* *Am. J. hum. Genet.* **28**, 280 (1976).

The Pope and the ontogeny of persons

John Godfrey

In his recent book *Crossing the Threshold of Hope*, Pope John Paul II airs his views on human reproduction. The pity is that he ignores most of modern genetics and embryology.

"THE concept of 'person' is not only a marvellous theory; it is at the centre of the human ethos. . . in this field more than in any other, collaboration among pastors, biologists and physicians is indispensable". Thus the Pope, in his new book *Crossing the Threshold of Hope*. In a discussion of St Thomas Aquinas, he writes: "It is not good that his thought has been set aside in the post-conciliar period; he continues, in fact, to be the master of philosophical and theological universalism". Why, then, does the Pope dispense with both modern biological knowledge and Aquinas's teachings when dealing with the issues of human reproduction?

The fundamental biological and ethical issue he raises is the origin of an individual person during life before birth. There is no moment when human life starts. It is clear that both the egg and the sperm are alive, and that their life is human, and that life is continuous from one generation to the next. Yet life, at this stage, is not yet that of a person.

The Pope accepts the misconception that there is an instant when fertilization happens. The process of fertilization is complex, taking about two days. It can involve more than one sperm nucleus, at least for a time. During this period genetic identity is yet to be established. There is a rapid and irreversible change at the activation of the egg. Activation, though rapid, is not instantaneous. Its very speed is probably an evolutionary adaptation of eggs to inhibit fertilization by many sperms. The activation that blocks polyspermy has two distinct phases: the first is quick but incomplete, the second slower but complete. A more severe difficulty is that activation, essential though it is to further development, precedes the chromosomal events that establish the genetic identity of the zygote. So a new individual cannot have her or his origin at activation. For about the first four days, all the genetically determined properties of the fertilized egg are maternal. Only after this do paternal genes begin to act, with gene expression characterizing the new individual.

Even then, stable individuality is some way off. Genetically different cells may still be exchanged between twins. Until a couple of weeks after fertilization a single embryo may still divide to produce identical twins. The component parts of what now becomes an unambiguous individual gradually develop. Crucial to the function of the brain is the laying down of insula-

tion along the nerve cells. This starts late in pregnancy but is not complete until after birth.

The picture of the nascent human that emerges from modern embryology is one of seamless change. There is no biological discontinuity that would be a candidate for a moment of unique philosophical significance, whether the advent of the individual person or the infusion of the soul. Thus, we should envisage human persons coming into being by continuous progression during ontogeny, as they did during phylogeny; as each bright day is imperceptibly created during dawn. So the rights of, and our duty towards, the unborn must similarly grow gradually.

It is understandable that ethical thinking should have been founded on obsolete biology. Roman Catholic doctrine is that the soul is infused during pregnancy at a particular, but unspecified, time. The present teaching of the Pope, and his church, short of doctrine, is that there is a serious possibility that this time is that of conception. Assuming that this is so avoids the danger of committing unintended homicide. I am not sure if 'conception' is meant to mean implantation, or the fertilization that precedes it. The distinction is not trivial. Many fertilized eggs fail to implant quite naturally. Is each of these to be considered the death of a person, even though most pass away unnoticed? Other eggs may not lead to a pregnancy because the woman is using an intrauterine device as a contraceptive. However, this acts to prevent implantation in the uterus, not fertilization. Aquinas, for long the received philosopher of the Catholic church, based his views on the observations and opinions of Aristotle. He believed that animation (ensoulment) occurred in three stages that were completed around 40 days after conception in the case of males, or 90 days in the case of females. Apart from the exotic error about the supposed difference between the sexes, based on the misleading evidence available to him, Aquinas's thoughts have, as so often, a disturbingly modern ring. His view that ensoulment is progressive is more nearly in tune with our present knowledge of biology than the instantaneous picture of events that the Pope adopts. Until the late nineteenth century, Christian tradition, following Aquinas, tended to grade the protection to the incipient person according to the stages of its development.

The Vatican has serious problems on

the issues associated with human reproduction, with both the laity and the priesthood. The bare facts of demography show that many Catholics do not practise what is preached by their church. The gulf between Rome and Italy on contraception and abortion is clear from social surveys and from the permissive legislation that was brought in by the Italian parliament. This might not be so serious if leading thinkers within the church were not so divided. The protest links liberal thinking on reproduction with objection to what is seen as authoritarian reaction to it by the Vatican.

Not only has the Catholic church recently taught that all stages of human development are morally equivalent but it has also adopted an unbending moral legalism in the application of the consequences of this view. This rigidity is not in keeping with important strands of traditional teaching. Aquinas held that fairness or good sense (*Epikeia*) should be used to recognize when human law might be inappropriate in a particular case. In his words: "Laws are made for human actions. But such actions are individual and concrete situations, and they are infinitely variable" and "The law should not be followed when to do so would be wrong". His theory of exceptions is in contrast to current unbending interpretation of his theory of natural law. The application of *Epikeia* enhances justice, revealing glimpses of a higher 'law beyond the law'. It has been possible to ignore his humane view because he wrote so briefly of it in his *Summa Theologiae*, probably because of the way he fitted it into the symmetrical formal design of his work.

In a finite world with limited resources, the biological issues with which the Pope deals affect everyone. The recent United Nations population conference made little progress partly owing to pressure from the Vatican (see *Nature* 371, 185; 1994). Pope John Paul should, first, consider afresh Aquinas's writing on fairness, and on the unique nature of people and their circumstances; and, second, look at the light that modern genetics and embryology sheds on the concept of a person and of her or his origin. He could then properly advise people on how best to lead their sexual and reproductive lives. □

John Godfrey, associate of the Centre for Human Ecology, University of Edinburgh, is at 41 Lawford Road, London NW5 2LG, UK.

The wonders of the microlaser

A group at MIT has managed to build and operate a laser that functions as such even when the average number of photons in the optical cavity is less than one.

EVERYBODY, of course, now knows how a laser functions. Its essentials are an optical cavity, which can be traversed repeatedly by radiation of an appropriate frequency and in a predetermined mode. The optical performance of the instrument depends on the degree to which modes (of optical vibration within the cavity) other than that desired are suppressed by the geometry of the cavity and on the avoidance of energy loss, as in reflection at the walls. The more effectively the losses are constrained, the more often a coherent light beam will traverse the cavity and, so, the more accurately determined will be the frequency.

But the optical cavity is only half of a laser. The other indispensable part is an original source of radiation whose frequency is approximately that for which the cavity is designed which, typically, is a supply of atoms in an excited state. Ideally, de-excitation of these atoms generates photons with a fixed frequency which, in traversing the cavity, stimulate the emission of coherent photons from other (excited) atoms, thereby generating the coherent oscillation of radiation within the cavity, part of which can be bled off into a usable laser beam. The intensity of the beam is determined by the rate at which excited atoms can be supplied, but its optical properties are largely determined by the quality of the cavity. There are obvious trade-offs between usable intensity and the quality of the beam.

That is a simplified version of what all the textbooks now say. Much of the interest in the development of lasers in the past three decades is the diversity of the uses to which they can be put. At one extreme, they are high-intensity sources of radiation much drooled over by those concerned with ballistic missile defences; at the other, they are frequency standards capable of remarkable accuracy. For applications of the first kind, design effort is lavished on the radiation source; for those of the second kind, the cavity naturally gets the attention.

In the second connection, the 'one-atom laser' now makes its appearance. At first sight, such a device sounds a contradiction in terms. If, as is seriously intended, there is only a single excited atom in the optical cavity, how can excited atoms act in concert to reinforce a radiation beam oscillating within the cavity? But a little thought will show that the contradiction is not as glaring as it seems. If the performance of the cavity part of the laser is almost ideal, the oscillating radiation package within the cavity will retain the properties it is designed to have

with only an occasional reminder, in the form of an authentic photon from a single atom, of what the frequency is meant to be.

At this stage, the Massachusetts Institute of Technology (MIT) group responsible for the new laser, called a 'microlaser', are less concerned with the development of frequency standards than with testing the elementary predictions of quantum electrodynamics (QED) that, in an accurately designed optical cavity, only modes of vibration determined by the geometry will be accessible to photons. The authors of this new development are Kyungwon An, James J. Childs, Ramachandra R. Dasari and Michael S. Feld (*Phys. Rev. Lett.* **73**, 3375–3378; 1994). Their first working version of a microlaser is clearly only a model for better things to come.

How does it work? The first need is for a 'supercavity', in this case fashioned from two convex mirror surfaces (with radius of curvature of 10 cm) separated by roughly 1 mm and held apart by a piezoelectric spacer that can be used to tune the resonant frequency of the cavity in real time, on a timescale of a few hundred nanoseconds.

The atoms occasionally injected into the cavity are atoms of ^{138}Ba (the naturally predominant isotope of barium) prepared as an atomic beam (by heating barium metal in a furnace with a narrow aperture) where they can be conveniently put into an excited state by means of a pumping laser. The effectiveness of the arrangement hangs crucially on the precision of the frequency of the laser that excites ^{138}Ba atoms before they reach the cavity. The MIT group says that, by means of an optical frequency-locking arrangement, they can control the frequency of the exciting laser (at least for periods of the order of milliseconds) to within 40 kHz which, for radiation with a wavelength of 791 nm, amounts to a variation of 1 part in 10^{10} in frequency. This precision is needed to ensure a steady supply of excited ^{138}Ba atoms.

At this early stage, the interest of the microlaser now described is largely that it can be built and operated more or less as intended. It is a technical *tour de force* whose importance is entirely belied by the modest language in which it is described. Astonishingly, the group responsible has found incontrovertible evidence of laser function when the atomic beam is so stopped down that the average number of barium atoms in the cavity is only 0.1. With a more plentiful supply of barium, amounting to an average presence in the cavity of 0.71 at-

oms, the laser signal is actually stronger and sharper than calculated.

There seems no doubt that the instrument functions as a means of exciting in the cavity only the particular mode of oscillation for which it is devised, and essentially at the frequency of the 791 nm barium transition. Indeed, the authors show a plot of the number of photons in the oscillating cavity as a function of the average number of barium atoms it contains. With an average of 1 atom within the cavity, the number of oscillating photons is merely 10. With half an atom (on the average) in the cavity, there is (on the average) just one oscillating photon in the cavity.

The authors write laconically of the improvements of the experiment they now plan. The atomic beam they have used consists of the raw output of the furnace in which barium is heated, which means that the velocity of the atoms follows a Maxwell distribution; velocity selection must give better control of the initial state of the excited ^{138}Ba atoms. Even better mirrors will yield a superior supercavity, which in turn should mean that fewer ^{138}Ba atoms will excite single photons in predetermined states of oscillation.

The immediate objective is, for the time being, pure physics: to provide direct tests of what QED has to say about the interactions between atoms capable of radiating photons and the quantum states of the electrodynamic vacuum. While the success of QED is now so celebrated that doubts of its correctness are non-existent, it will be good to see it all confirmed. Indeed, there will no doubt be some brave soul who will seek to use such techniques for telling whether photons occupying the same quantum state in a resonant optical cavity interact with each other, subtly changing the average energy of occupancy.

There are other possibilities, not the least of which is the long-standing goal of telling whether it is possible to circumvent the limitations of quantum measurement for photons, perhaps even of measuring the number of photons in a quantum state in a resonating oscillator without changing the number of photons in that state. Several ingenious ways of doing this have been proposed, but none of them has so far been adequately tested.

But these are still early days. As with other novel devices, there is no telling what, in the long run, will become of the microlaser. Meanwhile, it is a piece of sheer cleverness that deserves applause.

John Maddox

Virological mayhem

Simon Wain-Hobson

ONCE again we're back with HIV and AIDS, this time in the form of the data analyses by Wei *et al.*¹ and Ho *et al.*² that appear on pages 117 and 123 of this issue. The two groups have looked at the relationship between production of HIV-1 virions and turnover of CD4 T lymphocytes, the preferred target for HIV, in patients undergoing treatment with various drugs. The picture that emerges is of a titanic struggle between the virus and the immune system.

The background to this work lies in the rediscovery of HIV in lymphoid organs³⁻⁵ and the realization that there is far more virus and infected cells around than hitherto thought⁶⁻⁸. Longitudinal studies have shown that the overall viral load, of both free virus and that within infected cells, increases slowly but inexorably, generally over a period of months to years.

Wei *et al.* and Ho *et al.* show that treatment of patients with a number of anti-HIV drugs results in a rapid reduction of free virus in the plasma within a few days. This reduction is accompanied by an increase in the number of circulating CD4 T cells. A dark side of treatment is the rapid appearance of drug-resistant mutants; a single base change is often sufficient to confer resistance. Whether the target is the viral polymerase or protease, resistant virus may be detectable within a week⁹.

Similar data have been around for a year or two and are familiar to the specialist. So what's new? It is precisely the next step, so obvious in retrospect, that has now been made. Wei *et al.* and Ho *et al.* realized that such viral ups and downs provide us with the first glimpse of the dynamics of HIV production. By teaming up with mathematicians they were able to extract the pertinent information; the concordance of their data is remarkable.

Both groups treated mainly symptomatic patients with one of several different anti-HIV drugs. Extrapolating from the changes in the blood they could show that between 10⁸ and 10⁹ virions — the viral infectious units — were being cleared every day, equivalent to 30% of the total. As the levels of blood-borne viraemia show only gradual changes over many months, the daily rates of virus destruction and production must be comparable; that is, HIV production must be around 10⁸–10⁹ per day. The values differed little between patients. Wei *et al.* were even able to show that the emergence rate of drug-resistant virus ($t_{1/2}$ ~2 days) was virtually identical to clearance rate at the start of therapy.

The concomitant recovery of CD4 cells following drug treatment was modelled

by both groups, who assumed that the dynamics of blood lymphocytes (which represent some 2% of the total) reflected events throughout the body. Accordingly an average of 2×10⁹ cells per day (range 0.1–7×10⁹) must have been produced following drug treatment. Once again the rate of destruction of CD4 T cells could be inferred from the fact that changes in the absolute number of CD4 T cells is long (months to years) compared to the time-scales involved here (days to weeks). The relatively small net losses of CD4 T cells, 20–200×10⁶ per day, a figure derived from longitudinal studies of patients, represent the difference of two larger numbers. There is nothing surprising about rapid turnover in the immune system^{10,11}. In the best studied system, the mouse, half of most B lymphocytes are renewed within ten days¹⁰.

What is responsible for such efficient virus clearance and CD4 T-cell loss? As an intrinsic cytopathic effect of the virus is no longer credible^{12,13}, the immune system is the obvious answer. Antibodies specific for HIV, complement and macrophages could cope with the virus. Most of the infected CD4 T cells are to be found deep within the secondary lymphoid organs³⁻⁵. Yet infection results in the infiltration of HIV-specific cytotoxic T lymphocytes (CTL) or 'killer' cells^{14,15}. The CTL are consummate brokers of viral infections and are highly efficient in clearing viruses¹⁶. So long as a few virus proteins are being produced, the infected cell is marked out as foreign and ripe for destruction, often before the cell has had time to make completed virions.

The finding that more cells turn over than free virus would at first sight seem counter-intuitive, particularly as CD4 T-cell loss is a consequence of viral infection. This shows that total virus is underestimated by the two new studies, the virus in the periphery probably representing the excess production. But it is exactly what would be expected if CTL were stopping the vast majority of infected cells from producing progeny, leaving only a very small proportion able to produce virus. Indeed the tempo of virus sequence divergence requires such a result¹³. Note that CTL succeed by killing the virus-infected cell, so that the price to pay for viral clearance is a small part of the host. This is not a problem in an acute viral infection which is over in days, leaving ample time to recover the lost cells.

That billions of virions and infected cells can be destroyed every day vividly illustrates the very hostile environment created by the immune system — the meanest of streets are nothing by compari-

son. HIV counters by massive force of numbers. So long as a few progeny survive to continue replication and transmission then, from the point of view of the virus, that is a success story despite the billions destroyed. That the surface is, for many years, relatively calm is a tribute to the heroic efforts of the immune system. With so much virus and so many infected cells around, the AIDS-without-HIV hypotheses can definitely be sidelined.

Are there any practical implications of these new findings? Given that the virus is replicating 24 hours a day and from day one, antiviral treatment is called for at all stages of disease. Any means to reduce viral spread will ultimately be beneficial. Yet an asymptomatic patient can harbour at least 10⁶ genetically distinct variants of HIV, and for an AIDS patient the figure is more than 10⁸, among which one may find drug-resistant mutations even in the absence of therapy¹⁷. Against this background, monotherapy cannot succeed. Only combinations of drugs have the potential to outgun the virus. The recovery of the CD4 count, even in symptomatic infection, implies that immune reconstitution may be possible as long as replication of the virus can be held in check.

The new analyses are most welcome because they clarify the picture considerably. They show that HIV is behaving more and more like a virus, without frills or special effects. It is unique and subtle, but a virus nonetheless.

Finally, at a time when there is discussion as how best to channel resources for AIDS research¹⁸, it is worth noting that the main elements of our current picture of AIDS have come from old HIV hands rather than from insights into other viral or microbial lifestyles. True, there has been much poor AIDS research, many trivial theories and — yes — an appreciation of the immunopathology¹⁶ was long overdue. But the analyses of Wei *et al.* and Ho *et al.* show that there is a strong signal in AIDS research despite all the noise. Things are finally getting interesting. □

Simon Wain-Hobson is at the Institut Pasteur, 28 rue du Dr Roux, 75724 Paris Cedex 15, France.

1. Wei, X. *et al.* *Nature* **373**, 117–122 (1995).
2. Ho, D. D. *et al.* *Nature* **373**, 123–126 (1995).
3. Racz, P. *et al.* *Progr. Allergy* **37**, 81–181 (1986).
4. Pantaleo, G. *et al.* *Nature* **362**, 355–358 (1993).
5. Embretson, J. *et al.* *Nature* **362**, 359–362 (1993).
6. Platak, M. *et al.* *Science* **259**, 1749–1754 (1993).
7. Simmonds, P. *et al.* *J. Virol.* **64**, 864–872 (1990).
8. Connor, R. I., Mohri, H., Cao, Y. & Ho, D. D. *J. Virol.* **67**, 1772–1777 (1993).
9. Richman, D. D. *et al.* *J. Virol.* **68**, 1660–1666 (1994).
10. Freitas, A. A. & Rocha, B. B. *Immun. Today* **14**, 25 (1993).
11. Michie, C. A., McLean, A., Alcock, C. & Beverley, P. C. L. *Nature* **360**, 264–265 (1992).
12. Phillips, A. N. *et al.* *Nature* **367**, 124 (1994).
13. Wain-Hobson, S. *Nature* **366**, 22 (1993).
14. Tenner-Racz, K. *et al.* *Am. J. Path.* **142**, 1750 (1993).
15. Cheynier, R. *et al.* *Cell* **78**, 373–387 (1994).
16. Zinkernagel, R. M. & Hengartner, H. *Immun. Today* **15**, 262–268 (1994).
17. Najera, I. *et al.* *J. Virol.* **69**, 23–31 (1995).
18. Fields, B. *Nature* **369**, 95–96 (1994).

Of dusty tori and black holes

Richard Barvainis

It has been almost ten years now since a powerful new idea burst onto the scene of extragalactic astronomy¹. The idea, indirectly inferred from observational data, unified the two main spectroscopic categories of active galactic nuclei (AGNs) by positing that their apparent differences are due mainly to differing orientations with respect to the Earth. The hypothesis that a dusty molecular torus was the agent responsible for these orientation effects has since gathered much indirect support, but little direct observational confirmation. Now it appears that direct evidence may finally be at hand, in the form of two reports — one on page 127 of this issue², the other³ to be published next month — of water masers in the nucleus of the nearby active galaxy NGC4258. This work also provides what may constitute the strongest case yet for a super-massive black hole in a galactic nucleus.

On its introduction, the torus model immediately became popular, and has continued to gain steam, because of its compelling if not unique ability to explain the differences between the so-called type 1 and type 2 AGNs within a unified picture. In this picture the two types are intrinsically identical — or at least very similar — in nature. Observationally, the type 1 AGNs show extremely broad atomic emission lines in visible light, together with narrow lines, whereas the type 2s show only narrow lines. In the unified model a dusty molecular torus, or doughnut, surrounds the nucleus in both types of object, and both types do indeed produce both broad and narrow lines. But the broad-line region lies inside the torus and is blocked from view for some angles, while the narrow lines originate outside and are always seen. An object's classification is thus determined by whether its orientation is such that our view to the centre is clear or blocked. In this sense the torus has been seen mainly in shadow rather than in light.

The results of Miyoshi *et al.*² and Greenhill *et al.*³ give us a fascinating new peek at the torus itself, as probed by maser emission from gaseous water molecules. A maser (the 'm' is for microwave), like a laser, has the property of amplifying, sometimes by huge factors, background radiation or even its own internally pro-

duced emission. In astrophysical settings water masers are formed in relatively dense gas clouds, obviously containing water (and other) molecules and also often dust. In our Galaxy they occur in molecular clouds that are forming new stars or in the dusty, molecule-rich atmospheres of late-type stars. Water masers have been seen in other galaxies as well.

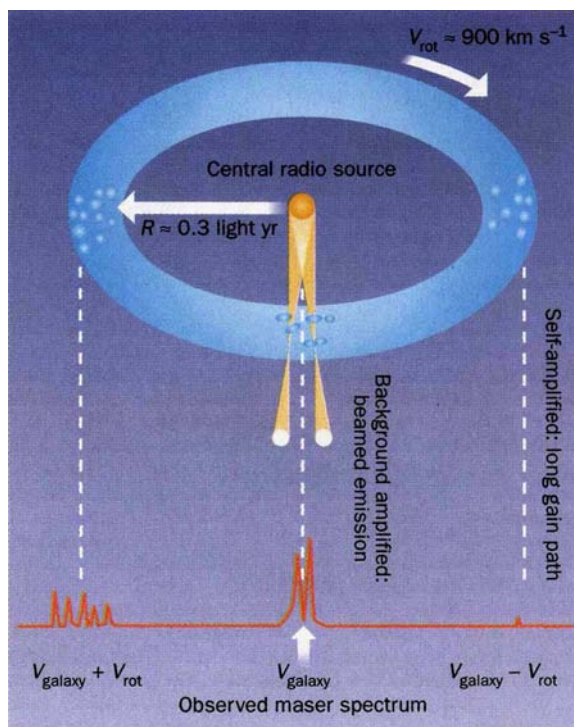


Illustration of the maser model, adapted from ref. 3. The masers arise in an annulus rotating at about 900 km s^{-1} , possibly the dusty molecular torus inferred to exist in active galaxies from other evidence. The maser spectrum, at short radio wavelengths near 1.3 cm , is sketched along the bottom. The clouds responsible for the systemic masers, in the centre of the spectrum, amplify the central radio source and become visible when they pass in front of it. The outlier features, at $\pm 900 \text{ km s}^{-1}$, have longer path lengths through a number of clouds and amplify their own intrinsic radiation. The orbital rotation of the torus about the central source should cause a slow drift of the clouds from west to east (right to left) in the case of the systemic masers. New features will appear as new clouds drift in front of the central source, and old features will disappear as clouds drift out. The outliers should appear nearly stationary, although perhaps variable with time as clouds orbit into and out of the tangent points.

One of these other galaxies is NGC4258, a nearby spiral at a distance of 21 million light years which is bright enough to have been noted by the famous eighteenth-century cataloguer of galaxies, Charles Messier, leading to its original designation as M106. It shows signs of activity in its nucleus, although relatively mild by AGN standards. Even though NGC4258 is considered a close neighbour, it is still so far away that the masers must

be extremely powerful to be seen at all. Because of their high luminosities, some of the masers in external galaxies have been called megamasers, and as they are much more powerful than masers in the Milky Way, their origin has been unclear.

The new work uses the radioastronomical technique of very-long-baseline interferometry (VLBI) to reveal the megamaser properties. The advantage of VLBI, which uses many telescopes spread across a continent or even around the world, is angular resolution — the ability to discern very fine details in the distribu-

tion of the masers on the sky.

Three observed characteristics are important to interpreting the VLBI results. The first is the general position of the masers, which is coincident with the nucleus of the galaxy. The second is that the individual maser spots are strung out along a line, as projected onto the sky, suggesting that they come from a flattened distribution of gas clouds. The third is the variation of velocity of the masers with position along the line: the velocity marches from negative to positive (relative to the centre) as one moves from west to east, meaning that on one side the masers are coming towards us, and on the other they are moving away.

The simplest explanation for these facts is that the masers originate in an edge-on, rotating annulus of molecular clouds situated in the galactic nucleus — the long-sought dusty molecular torus. One minor complication, which does not detract from this picture and in fact enhances it, is that there are three sets of maser features in NGC4258 widely separated in velocity. There is a 'systemic' set, at about the same velocity as the rest of the galaxy, and two 'outlier' sets symmetrically offset at plus and minus 900 km s^{-1} relative to the systemic set. A compelling model²⁻⁴ can be constructed assuming that both the systemic masers and the outliers come from a torus, albeit a rather flattened one.

In this picture, the systemic maser clouds are those that lie

along the line of sight to the central nucleus and amplify its radio emission, while the outliers come from the tangent points of the torus (see figure). The path lengths through the masing clouds at the tangent points are long enough so that the maser emission is self-amplified, and needs no background source. The model describes a torus with an inner radius of about 0.3 light years in precise keplerian orbit around an interior mass of about 40

million solar masses. A black hole of this size is quite respectable, and is in accord with the few available estimates of central masses in other active galaxies⁵⁻⁷.

The derived lower limit to the central mass density, more than 100 million solar masses per cubic light year, is higher than has been derived for any other galaxy. It seems likely that this mass is indeed in the form of a supermassive black hole, as, according to Miyoshi *et al.*, the alternative of a dense star cluster would be disrupted through stellar collisions on a timescale short compared with the galaxy's lifetime. If such a cluster were composed of Sun-like stars, their average separation would be only about 130 astronomical units, or a bit more than one and a half times the diameter of Pluto's orbit. The view of the night sky from a planet orbiting one of these stars would be spectacular, to put it mildly, but it could not last.

So what is left to be done? Additional VLBI observations are crucial to tying these ideas together and confirming the model predictions. Regular, frequent, high-resolution monitoring will now be possible using the new United States national VLBI facility, the Very Long Baseline Array (VLBA), which was the instrument used by Miyoshi *et al.* This is just the sort of project the VLBA was designed for, and measurements of any motions of the masers should be forthcoming within a couple of years. The model predicts that individual systemic masers should drift from west to east at a well-defined rate, and that the outliers should be stationary. There is already evidence for this from single-antenna spectral measurements, which show regular velocity shifts over time in the systemic masers but none in the outliers^{8,9}.

So the torus not only unifies two major categories of AGNs, and is interesting in its own right, but may also provide a rare opportunity to 'weigh' the central mass of an active galaxy. Direct evidence for the dusty molecular torus model has taken some ten years to come forward, and will take a few more to confirm definitively. But this is short as astronomical timescales go. Long live the torus! □

Richard Barvainis is at the MIT Haystack Observatory, Route 40, Westford, Massachusetts 01886, USA.

Under the ringed basins

H. J. Melosh

EARLY in the history of the Solar System a fusillade of cosmic debris scarred the surfaces of the Moon and terrestrial planets. The largest impacts on the Moon not only dug deep craters into its crustal rocks, but raised peculiar mountain chains encircling the impact sites. The origin of these mountain rings has been hotly debated since they were first recognized¹ around the Orientale basin in 1962.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Magellan image of the multi-ring basin Lise Meitner, 150 km in diameter. This venusian basin is surrounded by several steep inward-facing scarps that may be the surface traces of inward-dipping normal faults.

Although multi-ring basins like Orientale have now been identified on Venus², as shown above, as well as the Moon (there is some controversy about whether such basins occur on Mars and Mercury), even the best spacecraft images tell us little about what lies below the rings. So it is of great interest that Spray and Thompson, on page 130 of this issue³, report the discovery of what may be the deeply eroded traces of rings surrounding the 1.85-billion-year-old Sudbury impact structure in Canada. Their discovery suggests that the rings are the surface expressions of deep-seated faults, and thus rules out several proposed explanations for how the rings form.

Spray and Thompson argue that the rings left their marks in the form of pseudotachylite veins and dykes. Pseudotachylite is a dark, fine-grained rock that forms by frictional melting when adjacent rock masses slide relative to one another deep underground. Although it was first described in association with the Vredefort structure in South Africa, another probable impact site, pseudotachylite is not unique to impacts: rapid tectonic movements, probably associated with earthquakes, can also produce thin veins of melted rock deep in fault zones⁴.

What is unique to Sudbury and

Vredefort is the sheer size of the dykes — up to 100 m thick at Sudbury. Frictional melts of this thickness require large displacements to supply the heat needed for melting. At a depth of 5 km below the surface at least 1 km of motion between crustal blocks is needed to melt a 100-m-thick dyke³, and this displacement must have occurred more rapidly than the heat could have been conducted away from the

Washington Univ./NASA

sliding surface. The pseudotachylite dykes at Sudbury thus indicate very large, rapid displacements on a deeply buried fault surface.

The other significant fact about the Sudbury pseudotachylite dykes is that they are not scattered at random around the impact site. Very close to the edge of the former impact melt sheet (formally known as the Sudbury Igneous Complex) these dykes are pervasive and may have developed either during the passage of the shock waves from the impact, or as blocks surrounding the collapsing transient crater grated against one another.

However, a radial transect made by Spray and Thompson³ north from the centre of the complex shows several concentrations of dykes along what may have been crudely circular rings surrounding the melt sheet. Furthermore, the ratios of the diameters of adjacent rings seem to fall close to the approximate $\sqrt{2}$ ratio sometimes seen for the rings of extraterrestrial multi-ring basins.

What do these observations tell us about the formation of multi-ring basins? If Spray and Thompson's pseudotachylite dyke concentrations correlate with the steep scarps of basins such as Orientale on the Moon, or Meitner on Venus, then these scarps are almost certainly fault scarps with large displacements that developed very quickly, probably during collapse of the deep transient crater cavity. Unfortunately, the surface morphology of the basin is not preserved at Sudbury, but the $\sqrt{2}$ spacing supports such a correlation. Many planetary geologists have concluded independently that the observed scarps are the traces of faults.

If the rings are indeed fault scarps, a number of models for ring formation are ruled out. The nested crater hypothesis, for example, identifies the rings with the bevelled edges of uplifted lower crustal layers, and thus predicts the existence of

1. Antonucci, R. & Miller, J. *Astrophys. J.* **297**, 621–632 (1985).

2. Miyoshi, M. *et al. Nature* **373**, 127–129 (1995).

3. Greenhill, L. J. *et al. Astrophys. J.* (in the press).

4. Watson, W. D. & Wallin, B. K. *Astrophys. J.* **432**, L35–L38 (1994).

5. Gwinn, C. R. *et al. in Sub-arcsecond Radio Astronomy* (eds Davis, R. J. & Booth, R. S.) 331–332 (Cambridge Univ. Press, 1993).

6. Kormendy, J. in *Testing the AGN Paradigm* (eds Holt, S. S., Neff, S. G. & Urry, C. M.) 23–34 (AIP, New York, 1992).

7. Ford, H. C. *et al. Astrophys. J.* **435**, L27–L30 (1994).

8. Haschick, A. D., Baan, W. A. & Peng, E. W. *Astrophys. J.* **437**, L35–L38 (1994).

9. Greenhill, L. J. *et al. Astr. Astrophys.* (in the press).

concentric annuli of different composition beneath the rings. Such annuli are not seen at Sudbury. The Sudbury observations also rule out the old volcanic modification hypothesis, which likens ring formation to volcanic caldera collapse, although this hypothesis has long been discarded on other grounds. As the rings seem to have been defined by discrete faults, they also rule out the class of theories that attributes ring formation to broad waves arising as the debris filling the crater sloshes back and forth. These wave theories are, however, still viable for the formation of central peaks and peak rings in the interiors of large craters.

Although this glimpse below the surface of a multi-ring basin is of enormous value in understanding how rings form, it does not provide answers to all outstanding questions. The rocks preserved around Sudbury are too uniform to establish a stratigraphy, and so the real offsets across the putative ring faults cannot be determined. Similarly, details of the timing of crater excavation and fault slip cannot be tied down precisely.

Short of observing a fortuitous event

(like the crash of comet Shoemaker–Levy 9) that produces a multi-ring basin on a terrestrial planet, future researchers will have to continue to glean bits of evidence from the rocks around large impact structures on the Earth, use information from the surface morphology of extraterrestrial craters and exploit numerical models of crater excavation and collapse. The result of such research will be a better understanding of how large crustal masses respond to applied loads. Ring formation is evidently a deep-seated process and, if properly understood, may be used to probe conditions deep beneath the surfaces of the planets on which multi-ring basins form. □

H. J. Melosh is at the Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA.

1. Hartmann, W. K. & Kuiper, G. P. *Commun. Lunar planet. Lab. Univ. Arizona* **1**, 51–66 (1962).
2. Alexopoulos, J. S. & McKinnon, W. B. *Icarus* **100**, 347–363 (1992).
3. Spray, J. G. & Thompson, L. M. *Nature* **373**, 130–132 (1995).
4. Sibson, R. H. A. *Rev. Earth planet. Sci.* **14**, 149–175 (1986).

MOLECULAR RECOGNITION

Crystallographic antibodies

Michael Groß

THE growth of a snail's shell, the sudden freezing of frogs (which survive) and the swelling of a joint in an attack of gout all involve specific interactions between proteins and the surfaces or nuclei of small-molecule crystals. A remarkable — and possibly medically relevant — example demonstrating the specificity of this kind of process is described by L. Addadi and colleagues in the latest issue of *Journal of Molecular Recognition*¹. They show that antibodies induced by different urate-related crystal forms specifically catalyse the formation of new crystals of the kind they were raised against.

The notion that crystals of a compound that is not immunogenic in solution can give rise to antibodies came up only three years ago, when Addadi's group at the Weizmann Institute in Israel discovered that immunoglobulin G (IgG, the principal serum antibody) from patients with gout can accelerate crystallization of monosodium urate monohydrate (MSUM)². Accumulation of microscopic crystals of MSUM in the synovial fluid of the joints (resulting from high concentrations of sodium urate in the blood) is exactly what makes people suffer from gout³. So it was suggested that the presence of these crystals in the patients' synovial fluid had induced the formation of specific anti-crystal IgGs, which in turn, acting as a kind of catalytic antibody, were

able to promote the nucleation of more crystals of the same kind².

The new work shows that the interaction is specific enough to differentiate between MSUM, magnesium diurate octahydrate and the uncharged structural analogon allopurinol (which is also used in gout therapy). Addadi and colleagues raised antisera against these three crystal forms in rabbits, and measured the time needed for crystals to appear in the presence of pre- and post-immune sera or no IgG at all for the matching antibody–antigen pairs as well as, in a cross-reactivity test, for the other combinations. Each of the three IgGs strongly enhanced the formation of the crystal form it was raised against, while cross-reactivity ranged from weak positive to inhibitory (see figure).

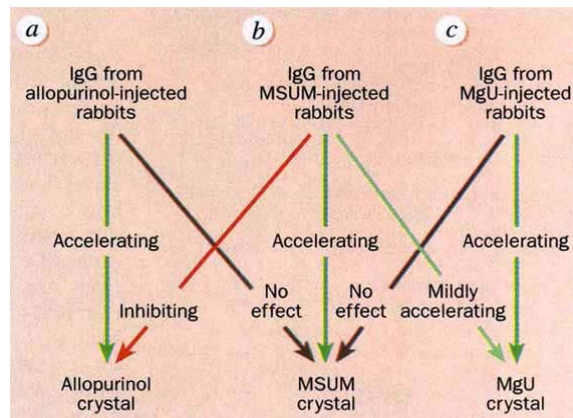
It has now been established that crystals are antigenic although they cannot be processed by the immune system in the usual way. This finding holds promise for the attempts to understand the inflammatory response in pathogenic crystallization or exposure to crystals, not only in gout

but also in other conditions such as silicosis and gall stones. The hypothesis that the IgGs catalyse crystal nucleation *in vitro* and *in vivo*, if borne out, would provide a neat explanation for the behaviour of the recurring attacks of gout, which become more frequent but less serious with time³.

To confirm this idea, however, the early events need to be studied more directly. The current work relies on the assumption that, under conditions where nucleation is extremely slow, effects on nucleation can be traced by observation of the appearance of macroscopic crystals. In a typical example of MSUM crystallization, crystals became observable after 7.5, 28.5 and 30.5 hours in the presence of post-immune, pre-immune and no IgG, respectively. Thereafter, crystallization needed less than one hour to proceed to completion. Although these results are convincing as evidence of the specificity of the antibody–crystal interaction, they provide only indirect indications as to what is going on during the first few minutes or hours.

In the case of protein crystallization — which is all too rarely compared with inorganic crystal growth⁴ — kinetic models are available to bridge the gap between the all-important early events and the more readily observable later stages^{5,6}. With small molecules, however, it may prove more difficult to characterize the nucleus as such, let alone its interaction with a macromolecule. It will be equally challenging to show that the rate-enhancing effect of IgGs is a catalysis of nucleation rather than, for instance, a consequence of kinetic partitioning between right and wrong aggregates (specific molecular chaperones are known to assist toxin crystallization in *Bacillus* species⁷) or of manipulation of oversaturation by changes of local concentrations.

Nucleation catalysts in biomineralization, for instance the highly acidic macromolecules in mollusc shells, often function



Specificity and cross-reactivity of IgGs raised against crystals of (a) allopurinol, (b) monosodium urate monohydrate and (c) magnesium diurate octahydrate (MgU). (From ref. 1.)

only when adsorbed to a solid phase (that is, when they are in the 'right place', where the cell needs the material to be crystallized), and inhibit crystallization when in solution⁸. Nonetheless, given the number of open questions in this field, that should not rule out the possibility of antibodies acting as nucleation catalysts in solution.

As to the structural details of the antibody-crystal recognition, the obvious step forward, namely the production of monoclonal antibodies, has already been performed — although with more difficulties than one would encounter with conventional antigens — and their characterization is in progress (L. Addadi, personal communication). In principle, the crystal structure of such an antibody might teach us quite a lot about gout, the mechanism of crystal nucleation and growth, antibody specificity and variability, and about molecular recognition between such apparently disparate partners as the regular array of a small molecule crystal and the highly asymmetrical and fractal structure of a globular protein.

This unequal match is no singular effect though. Last year, Addadi's group demonstrated recognition between enantiomorphic crystal surfaces and living cells by sorting morphologically achiral crystals of calcium tartrate tetrahydrate into the enantiomeric forms just by looking at which ones the cells preferred to settle on⁹. Obviously, nature has never drawn a sharp line between inorganic chemistry and biochemistry. Even if one does not believe Cairns-Smith's claim that life had a mineral precursor¹⁰, one has to acknowledge that molluscs have been using proteins to control crystal nucleation, growth rate and morphology for almost a billion years⁸. □

Michael Groß is in the Oxford Centre for Molecular Sciences, New Chemistry Laboratory, South Parks Road, Oxford OX1 3QT, UK.

1. Kam, M., Perl-Treves, D., Sfez, R. & Addadi, L. *J. Molec. Recogn.* **7**, 257–264 (1994).
2. Kam, M., Perl-Treves, D., Caspi, D. & Addadi, L. *FASEB J.* **6**, 2608–2613 (1992).
3. Dieppe, P. & Calvert, P. *Crystals and Joint Disease* (Chapman & Hall, London, 1983).
4. Rosenberger, F. *J. Cryst. Growth* **76**, 618–636 (1986).
5. Ataka, M. & Asai, M. *Biophys. J.* **58**, 807–811 (1990).
6. Groß, M. & Jaenicke, R. *Biophys. Chem.* **45**, 245–252 (1993).
7. Wu, D. & Federici, B. *J. Bact.* **175**, 5276–5280 (1993).
8. Addadi, L. & Weiner, S. *Angew. Chem. Int. Edn Engl.* **31**, 153–169 (1992).
9. Hanein, D., Geiger, B. & Addadi, L. *Science* **263**, 1413–1416 (1994).
10. Cairns-Smith, A. *Seven Clues to the Origin of Life — A Scientific Detective Story* (Cambridge Univ. Press, 1985).

Erratum

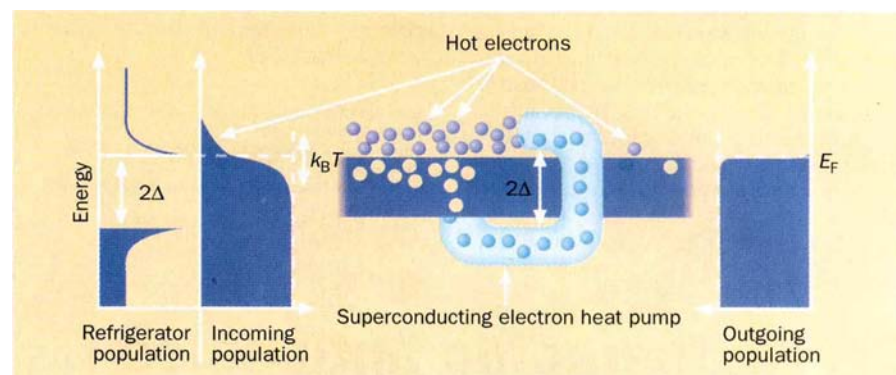
In last week's News and Views section, the author of "The signature of chaos" (page 16) was incorrectly given as Richard V. Jensen instead of Roderick V. Jensen.

A superconducting siphon

Sean Washburn

THE more accurate an electronic circuit must be, the cooler it must be. This is because the noise in the circuit results from thermally excited electrons hopping up and down among the available energy levels. The colder the electrons, the smaller the range of energy through which they are distributed and hence the more closely the circuit reads to the desired value. Following a proposal floated just over a year ago¹, M. Nahum *et al.*² have now demonstrated a new way of refrigerating the electrons in very small circuits.

Conventionally, electronic circuits are cooled by immersion in a refrigerant such as liquid helium or by a standard gas compression and expansion cycle similar to the one operating in home refrigerators. Both are cumbersome, and both rely on indirect cooling of the electrons by cooling the underlying crystal lattice and allowing the excess electron energy to be dissipated into lattice vibrations which are removed by the refrigerator. The fundamental difficulty with the conventional method is that the coupling between the



The electron refrigeration scheme proposed by Nahum *et al.*².

The electrons in any conductor can occupy a distribution of energy states. At absolute zero ($T=0$), the electrons fill all states up to the Fermi energy (E_F) and all states above that energy are completely empty. This is approximately the situation depicted as the outgoing population in the figure. Crudely speaking, one may say that as all of the electrons below the Fermi energy are 'surrounded' by electrons held in filled energy states, they are 'immobilized' and do not contribute to electric current in any circuit. Only the electrons at E_F have free states nearby to allow them to move and contribute to the electric current.

Any ambient thermal energy causes some fraction of the electrons to hop into higher energy states (leaving behind vacant states or 'holes' below E_F). As both the hot electrons and the electrons near the vacant states have neighbouring free states that allow them to move through the lattice, the energies of the electrically responsive electrons are smeared over the range of energies through which the electrons are excited, which is the thermal energy $k_B T$ (k_B is Boltzmann's constant). Instead of the contributions from only those electrons precisely at E_F , the circuit responds for energies within the range $\Delta E = E_F \pm k_B T$. Obviously in situations where the circuit must operate with very small ΔE , one desires a very low electron temperature.

electrons and the lattice is very poor at low temperature and vanishes rapidly (inversely with T^5) as the temperature decreases³. At very low temperatures ($T < 0.1$ K), the lattice temperature is nearly irrelevant to the electron temperature, which is easily elevated by even the tiniest circuit bias (for example, dissipation of 1 nW can increase the electron temperature by a significant amount). Ultimately, as one wishes to drive the electron temperature towards zero, cooling the lattice is fruitless.

The new refrigeration techniques^{1,2} rely on thermally isolating a small conductor and directly cooling the electrons. In fact (for the time being, at least) they must rely on conventional schemes to cool the lattice so that the electrons are not re-heated by lattice vibrations. Two different quantum mechanical schemes have been proposed for removing electrons above a certain energy and replenishing the Fermi sea with lower-energy electrons. These are quite different from brute-force cooling of the lattice or from more elegant means of cooling such as feedback cooling of the electron thermal fluctuations⁴.

The recent proposal for direct refrigeration of the electrons by Nahum *et al.*² employs a superconducting electrode in contact with a small metal conducting element. The defining feature of the superconductor is, of course, that there is a gap of width 2Δ in the available electron

states (the situation is depicted as the refrigerator population in the figure). Electrons in states below the gap are free to move 'frictionlessly' throughout the lattice, and there are very few electrons in the (non-superconducting) states above the gap. The voltage on the superconducting electrode is set so that the top of the gap is just above E_F . Electrons with energies greater than E_F tend to hop out into the empty states of the superconducting electrode, but the cooler electrons cannot hop into the gap (as there are no states there). Electrons from the states below the gap, and therefore necessarily considerably below E_F , can replace the hot electrons — or more precisely fill in the holes below E_F left by the hot electrons.

A (somewhat fanciful) diagram of the principle of the machine appears in the figure. The hot electrons appear as bubbles floating above the surface of the Fermi sea and the holes as empty regions in the sea. The superconducting electrode serves as a siphon to extract hot electrons (bubbles) and replenish them with cold electrons at $E_F - 2\Delta$, which fill in the holes in the Fermi sea. The replenishment process must be accomplished without pumping the heat right back into the conductor, but fortunately there is a process called Andre'ev scattering that permits the electrons to cross the boundary without carrying the heat back in^{5,6}. The authors not only proposed a scheme for reducing the electron temperature, they confirmed that it works in experiments on a bolometric circuit that requires the extreme energy precision afforded by very cold electrons.

The earlier direct refrigeration scheme¹ was proposed theoretically, but has not been tested. It accomplishes a similar 're-distribution' of the electron energies by a somewhat more complicated method. The individual quantum states of very small capacitance contacts ('quantum dots') are used as energy selective ports into and out of the conductor. The contacts must be adjusted so that energy levels in the inlet and exit contacts are carefully aligned near E_F . The alignment is tailored so that electrons are fed into the conductor through a single quantum level slightly below the Fermi energy and extracted through a level slightly above the Fermi level¹. As the electrons entering the conductor are at a low energy and those leaving are at a high energy, the net result is a lowering of the average electron temperature as above. The problems in this scheme are the difficulty of adjusting two separate rather delicate energy scales and that the extraction can remove only

electrons that are very near the energy level in the exit contact — all other electrons are impervious and sail past.

One should not expect to see such refrigeration schemes appear soon in common laboratory equipment. The superconducting electrode can remove only femtowatts of power, so it is not capable of cooling any but the tiniest of conductors and then only slightly: in ref. 2, a ten-micrometre long, half-micrometre wide strip of copper was cooled from 0.100 K

(the operating temperature of the conventional refrigerator, which cooled the lattice) down to 0.085 K. As both this temperature regime and circuit size regime are extreme, the refrigerator will remain a tool for high-resolution radiation and particle detectors. □

Sean Washburn is in the Department of Physics and Astronomy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA.

NEUROBIOLOGY

Thorny issues in neurons

Rodolfo Llinás

NEURONS can have thousands of synaptic inputs, each of which may have an excitatory or inhibitory effect of varying strength. So how are all of these inputs integrated, with the result that the neuron does or does not fire? As Eilers *et al.* point out in the introduction to their paper on page 155 of this issue¹, neuronal integration was at one time thought to be the simple summation of synaptic potentials. In the light of electrophysiological findings of the past two decades, that view has had to be modified. Nowadays, the complications engendered by dendritic geometry², the non-uniform resistive properties of the dendritic plasmalemma³ and the nonlinear electrical properties stemming from changeable membrane conductances⁴, as well as synaptic plasticity⁵, all have to be taken into account.



FIG. 1 Purkinje neurons have the largest dendritic tree in higher vertebrate brains. They receive excitatory synaptic input from the parallel fibres which contact their spines (not visible here) located at the thin, terminal dendritic branches. *a* and *b* in Fig. 2 correspond to a small area of dendrite (arrow).

Eilers *et al.* describe evidence for another level of complexity, one mediated by calcium at particular places in the dendritic tree. The neurons they worked with, Purkinje cells, occupy a privileged position in research of this sort. These cells have the largest dendritic arbour in the vertebrate brain (Fig. 1), and are elegantly isoplanar, making them especially appropriate for *in vitro* studies using brain slices. This dendritic tree is endowed with the most convergence (up to 200,000 parallel fibre inputs on to dendritic spines), and the most divergent synaptic organization (one spine per parallel fibre, except for the ascending axons of the granule cells⁴). These spines are a micrometre or so in diameter and are 'necked' (1–1.5 μm in length and 0.2 μm in diameter) to the Purkinje cell. The possible permutations for synaptic bombardment of this receptor surface are absolutely astronomical.

The question, of course, is how sensitive this cell is to each of its many spine inputs. Action potentials in a parallel fibre probably release one or two quanta of transmitter (and often none)⁶ onto its corresponding Purkinje cell spine. Is the potential generated by this input, when it occurs, large enough to activate voltage-dependent calcium currents in the receiving spine? This sensitivity will determine whether spines are simple sites for synaptic contact or whether they are more deeply involved in neuronal function.

Eilers *et al.* carried out experiments in which a small group of parallel fibres was electrically activated while the target Purkinje cell was recorded at the somatic level. The relationship between post-synaptic response amplitude and changes in local calcium concentration at the dendritic synaptic junctions was determined using the fluorescent dye calcium green under confocal microscopy. With convincing illustrations, the authors show that changes in calcium concentration may be restricted to spiny branchlets and spines. Indeed, their findings indicate that mini-

1. Edwards, H. L., Niu, Q. & de Lozanne, A. L. *Appl. Phys. Lett.* **63**, 1815 (1993).
2. Nahum, M., Eiles, T. M. & Martinis, J. M. *Appl. Phys. Lett.* **65**, 3123 (1994).
3. Roukes, M. L. *et al.* *Phys. Rev. Lett.* **55**, 422 (1985).
4. Washburn, S. *Nature* **337**, 597 (1989).
5. Andre'ev, A. F. *J. exp. theor. Phys.* **19**, 1228 (1964).
6. Washburn, S. *Nature* **361**, 587 (1993).

Forced wood

A CENTURY ago, engineers and carpenters could buy huge, cheap baulks of timber from the virgin forests of the world: kauri pine, yellow pine, teak and so on. Nowadays the virgin forests are largely despoiled, and wood from cultivated forests comes in smaller, more expensive chunks. Daedalus now has the answer. He is inventing factory forestry.

The most useful part of a tree, he says, is its trunk. The roots at the base provide it with a solution of mineral salts; the leaf-bearing branches provide it with photosynthesized glucose and remove excess water by transpiration. Given these simple services, the trunk will grow for ever. DREADCO engineers are therefore sawing young trees down, cutting off their branches, and sealing the stump ends to a hydraulic supply system. The base of each trunk is pressurized with mineral solution; the branch terminals are alternately sucked to extract water and pressurized with glucose solution (this mimics the diurnal variation of sap flow in a living tree). When the technique has been perfected, the trunk should grow as well as if it were still planted.

Indeed, it should grow much better. The hydraulic pressures and flow rates at DREADCO's command are much higher than anything found in nature. No longer limited by the photosynthetic capacity of its leaves and the intermittent supply of sunlight, the trunk will grow with amazing speed. The ultimate limit may be set by its need to dispose of metabolic heat. It could reach full natural size, and even more, in a matter of months.

Daedalus wants to install a second root pump at the 'top' of the tree. This should abolish its vertical taper and cause it to grow into a usefully uniform cylinder. Insulated from seasonal variations, the cylinder will expand smoothly and perfectly, without the weakening flaws, ring structure and grain of natural wood. Where grain is wanted (as in walnut veneers for furniture and dashboards), artificial 'seasons' could be imposed on the pumping regime: perhaps much faster than the natural ones, or with a more complex cycle.

A factory forest will be a compact, efficient sort of place. It will need copious water, and glucose from some cheap source such as sugar beet. Densely packed vertically or horizontally, and fed from a web of piping, its tree trunks will grow almost visibly until they reach full size, and can be replaced by fresh saplings. Cheap and perfect factory timber will bring back the golden age of timber engineering, and take the pressure off the threatened forests of today.

David Jones

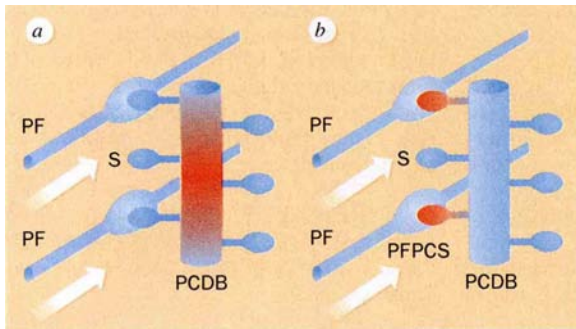


FIG. 2 Eilers *et al.*¹ demonstrate that parallel-fibre (PF) action potentials (arrows) which depolarize Purkinje-cell dendritic branchlets (PCDB) — through the parallel-fibre Purkinje cell synapse (PFPCS) — will activate calcium entry (red) into these fine processes, even when the voltage generated across the dendritic tree is subthreshold for initiation of an action potential in the Purkinje cell soma or dendrites. This voltage-dependent entry of calcium may be initiated at the dendritic branchlets (a) and spread to the spines (S), or at the spines (b) and spread to the branchlets. Although the time resolution of their instrumentation does not allow Eilers *et al.* to differentiate between these two possibilities, they support the first scheme; I favour the second. This point is seemingly one of detail, but is central for understanding neuronal integration and interpreting the functional mechanism for synaptic plasticity.

mal stimuli, capable of generating small depolarizations at the body of the cell, do not always produce an observed calcium response. As the stimulus is increased, calcium responses appear even before full somatic spikes (of sodium) or dendritic spikes (of calcium) are observed.

The authors report that calcium responses were limited to spiny branchlet segments and they did not observe single spines activated on their own. They conclude that spiny branchlets and associated spines can be activated as a whole, once enough parallel fibres are activated synchronously. They also suggest that activation of this branchlet-spine complex is initiated at the branchlet and is followed by spine invasion or, equally likely, that calcium entering the branchlet may diffuse passively into the spine (Fig. 2). The time resolution of the technique (<30 milliseconds) is too slow to establish this sequence, however, and so for the moment these ideas must remain conjectural.

Although the general result concerning subthreshold calcium conductances in the dendrites of Purkinje cells is in line with previous findings^{4,7}, the proposal that single spines are never activated alone is most significant and unexpected. Indeed, transmembrane currents produced by activation of a single parallel fibre (>40 picoamps)⁶ should generate voltages of 20 millivolts at the spine itself³ — which could be sufficient for P-channel activation⁸, given a dendritic resting potential of -60 millivolts, and so should be the first structure to respond; indeed, reaction products to calcium-channel antibodies are evident at dendritic spines⁹. This is a crucial question, as it indicates that either the present view of spine

impedances is wrong or that spines have no calcium channels. The final word on this issue awaits the development of technology with faster time resolution.

A different possibility comes from work involving activation of the ascending axon of parallel fibres, which contacts groups of spines on their upward path^{4,10}. In this case, branchlet response may be due to the simultaneous activation of such structures rather than the summation of independent activity from many parallel fibres.

A second notable point arising from the new findings¹ is that of long-term depression in Purkinje cells. It has been proposed that the synapse between a parallel fibre and a Purkinje cell is reduced in effectiveness when it is preceded by climbing fibre activation¹¹, and this has been put forward as the mechanism underlying learning in the cerebellum. The fact that parallel fibres themselves can activate voltage-dependent calcium conductances in dendrites implies that, unless the effect of climbing fibres on parallel fibres is not related to local dendritic calcium (as currently believed)⁵, long-term depression should also be generated by the inputs from parallel fibres alone¹². If this is true, the above proposal¹¹ is in trouble. Indeed, according to Eilers and colleagues' findings, the spines (where plasticity has been proposed to occur) might not be able to distinguish calcium entry dependent upon climbing fibres from that produced by branchlet activation from parallel fibres. Clearly, these are thorny issues and much remains to be learned. □

Rodolfo Llinás is in the Department of Physiology and Neuroscience, New York University Medical Center, 550 First Avenue, New York, New York 10016, USA.

1. Eilers, J., Augustine, G. J. & Konnerth, A. *Nature* **373**, 155–158 (1995).
2. Rall, W. & Segev, I. *Physiol. Rev.* **72**, 159–186 (1992).
3. Rapp, M., Segev, I. & Yarom, Y. *J. Physiol.* **474**, 101–118 (1994).
4. Llinás, R. & Sugimori, M. in *The Cerebellum Revisited* (eds Llinás, R. & Sotelo, C.) 167–181 (Springer, New York, 1992).
5. Konnerth, A. & Eilers, J. *Biomed. Res.* **15**, 73–77 (1994).
6. Barbour, B. *Neuron* **11**, 759–769 (1993).
7. Miyakawa, H., Lev-Ram, V., Lasser-Ross, N. & Ross, W. N. *J. Neurophysiol.* **68**, 1178–1189 (1992).
8. Usovich, M. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **88**, 7076–7080 (1991).
9. Hillman, D. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **88**, 7076–7080 (1991).
10. Jaeger, D. & Bower, J. M. *Exp Brain Res.* **100**, 200–214 (1994).
11. Albus, J. S. *Math. Biosci.* **10**, 25–61 (1971).
12. Batchelor, A. M. & Garthwaite, J. *Neuropharmacology* **32**, 11–20 (1993).

Exercise and brain neurotrophins

SIR — Physical activity has emerged as a predictor of high mental function in ageing¹. Exercise has been shown to affect several neurotransmitter systems^{2,3}, but it remains to be shown whether it can influence other key molecular systems which serve the maintenance and plasticity of the brain. Brain-derived neurotrophic factor (BDNF), a growth factor of the neurotrophin family, supports the function and survival of many neurons⁴⁻⁶, and may help protect neurons from free-radical damage⁷. Ample evidence indicates that production of BDNF in the brain is regulated by neuronal activity⁸⁻¹¹. Here we report evidence that physical exercise can increase BDNF gene expression in specific brain regions. These data open the encouraging possibility that exercise may increase the availability of trophic support, and thus resilience against insult, in certain neuronal populations.

Rats seem to find physical activity inherently rewarding; most will voluntarily run 1–2 km during the night, given access to a running wheel. Once acclimatized to running, each rat tends to maintain a characteristic activity level, running a similar total distance each night.

We used a nuclease protection assay to measure BDNF messenger RNA in distinct brain structures of adult rats following 0 (control), 2, 4 or 7 nights with free access to running wheels. BDNF mRNA in the hippocampus and caudal 1/3 of the

neocortex ($P = 0.05$) was significantly increased over control levels after 2 nights with exercise, and remained elevated for 7 nights.

We next examined BDNF mRNA

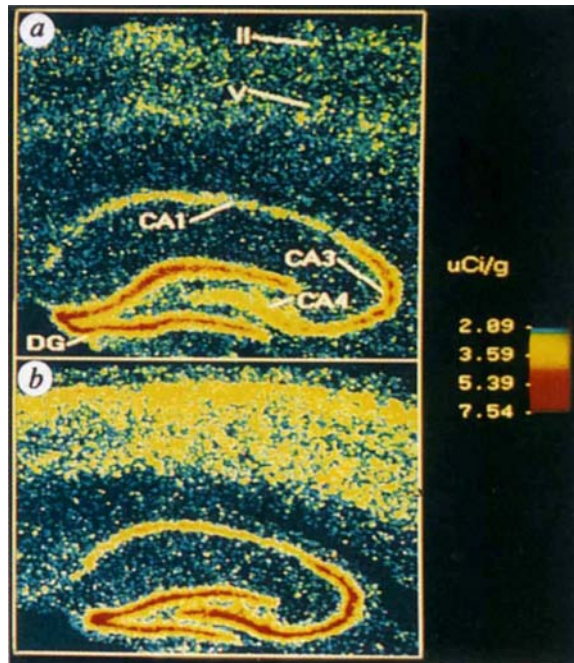


FIG. 1 Colour-enhanced image of autoradiograms from *in situ* BDNF mRNA hybridization in rat brain (sagittal plane). Increasing hybridization (ROD) is represented by green, yellow, then red. *a*, Control; *b*, 7 nights with exercise. Neocortical layers II and V, and hippocampal areas CA1–4 and dentate gyrus (DG), are identified. Methodological details can be obtained from the authors.

hybridization in specific cell populations within these structures via *in situ* hybridization (Fig. 1). Within the hippocampus, the pyramidal cells of Ammon's horn areas 1 and 4 (CA1, CA4) showed the largest upregulation, whereas CA3 and the dentate gyrus granular layer (DG) sustained smaller, later increases. In the caudal neocortex, labelling in layers II–III and V–VI rose progressively with

length of exposure to exercise.

Because individual rats in the study maintained different activity levels, we compared BDNF mRNA levels with distance run per night by each rat. We found a significant, positive correlation between mean distance run per night and BDNF mRNA in hippocampus (Fig. 2*a–c*) and caudal neocortex (data not shown) in all exercise groups. One rat in the 2-night group did not run, and had BDNF mRNA measurement (Fig. 2*a*) close to control levels.

Interestingly, the greatest effects of exercise on BDNF occurred in highly plastic, or changeable areas, responsive to environmental stimuli^{11–13}. This result supports previous suggestions that BDNF is involved in brain plasticity. A second relevant consideration is BDNF's potential involvement in neuronal survival and functional maintenance. For example, retrograde transport of BDNF from the hippocampus provides vital support to forebrain cholinergic neurons^{14,15}, a site of Alzheimer's disease and age-related degeneration. Physical activity could increase availability of BDNF to these cells by upregulating its expression in the hippocampus. Exercise-induced upregulation of BDNF could help increase the brain's resistance to damage and degeneration through BDNF's support of neuronal growth, function and survival.

S.A. Neeper^{††}

F. Gómez-Pinilla^{††}

J. Choi[†]

C. Cotman^{††§}

[†]Department of Psychobiology,

^{††}IRU in Brain Aging,

^{††}Department of Neurology,
University of California, Irvine,
California 92717,
USA

[§] To whom correspondence should be addressed.

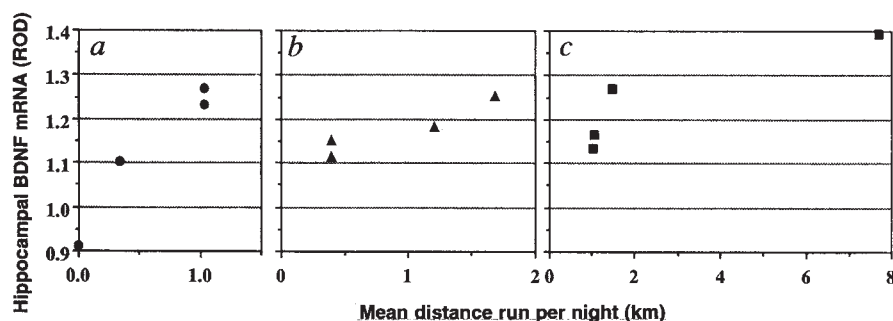


FIG. 2. Mean nightly distance run by each rat was significantly correlated with hippocampal BDNF mRNA (relative optical density measurements from the nuclease protection assay). *a*, Two-night ($r=0.967$, $P=0.05$); *b*, 4-night ($r=0.95$, $P=0.05$); and *c*, 7-night ($r=0.892$, $P=0.02$) groups. Wheel revolutions were monitored by computer (Ratrun, C. Hage Associates).

- Berkman, L.F. et al. *J. clin. Epidemiol.* **46**, 1129–1140 (1993).
- Fordyce, D.E. & Farrar, R.P. *Behav. Brain Res.* **43**, 115–123 (1991).
- MacRae, P.G. et al. *Neurosci. Lett.* **79**, 138–144 (1987).
- Altar, C.A. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 11347–11351 (1992).
- Alderson, R.F., Alterman, A.L., Barde, Y.A. & Lindsay, R.M. *Neuron* **5**, 297–306 (1990).
- Ghosh, A., Carnahan, J. & Greenberg, M. *Science* **263**, 1618–1623 (1994).
- Spina, M.B., Squinto, S.P., Miller, J., Lindsay, R.M. & Hyman, C. *J. Neurochem.* **59**, 99–106 (1992).
- Zafra, F., Castren, E., Thoenen, H. & Lindholm, D. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10037–10041 (1991).
- Okazawa, H., Murata, M., Watanabe, M., Kamei, M. & Kanazawa, I. *FEBS Lett.* **313**, 138–142 (1992).
- Isackson, P.J., Huntsman, M.M., Murray, K.D. & Gall, C.M. *Neuron* **6**, 937–948 (1991).
- Castren, E., Zafra, F., Thoenen, H. & Lindholm, D. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9444–9448 (1992).
- Diamond, M.C. et al. *Exp. Neurol.* **96**, 241–7 (1987).
- Volkmar, F.R. & Greenough, W.T. *Science* **176**, 1145–1147 (1972).
- Morse, J.K. et al. *J. Neurosci.* **13**, 4146–4156 (1993).
- Widmer, H.R., Knusel, B. & Hefti, F. *Neuroreport* **4**, 363–366 (1993).

Weathering and glacial cycles

SIR — Lovelock and Kump¹ present a simple model to explain how the terrestrial and marine biota might affect glacial–interglacial climate variation. In their general scheme, the optimal conditions for CO₂ uptake by biologically mediated weathering are the cool glacial periods. They thus suggest that during the glacial periods land vegetation mediates increased rates of chemical weathering uptake of CO₂ (relative to the interglacials), which could favour low CO₂ levels and help to stabilize glacial climates.

However, given that chemical weathering is strongly favoured by warm, moist conditions², there seems reason to believe that global weathering rates were higher during interglacial periods than during glacial periods. Generally, glacial maxima are very dry and cold on a global scale^{3,4}, and this would tend strongly to suppress weathering rates relative to the much warmer and moister interglacials. The accumulation of chemically unstable minerals in surface sediments during glacials should furthermore tend to promote a 'burst' of chemical weathering during the interglacial conditions that follow. Even though some parts of glacial phases do appear fairly moist on a global scale, at these times the cooler conditions should suppress weathering relative to the interglacial phase⁵. Thus there seems no strong basis for assuming that global CO₂ uptake by terrestrial weathering should be greater during glacial phases.

It seems that the conclusion of Lovelock and Kump¹, that the action of land vegetation as a catalyst for the CO₂–weathering sink is to 'pull' the Earth's climate system back down towards glacial levels, may be unjustified. Instead, variation in weathering rates seems likely to act as a damping agent on both CO₂ and climate on the glacial–interglacial timescale (dragging down CO₂ levels more strongly during interglacials, and relatively weakly during glacials), against a background of large oceanically driven CO₂ changes.

Changes in bulk terrestrial carbon storage in land ecosystems may have a greater influence on glacial–interglacial fluctuations in CO₂ levels. There are good reasons for thinking that this too would have acted as a negative feedback on CO₂ fluctuations between glacial and interglacial conditions³, paralleling the effects of weathering.

Jonathan Adams
School of Geography,
University of Oxford,
1 Mansfield Road,
Oxford OX1 3TB, UK

LOVELOCK AND KUMP REPLY — A conceptual paper such as ours¹ is undoubtedly subject to criticism, and, it was our hope that it would stimulate discussion. However, Adams' criticisms² are based on an obvious misreading of our paper. He claims that we propose "optimal conditions for CO₂ uptake by biologically mediated weathering are the cool glacial periods." This is a misstatement of our basic premise. In several places, including our Fig. 1, we point out that the ecological optimum for land plants is probably near a global average of 18 °C, certainly not the "cool glacial" state, but rather a state even warmer than today's.

Perhaps Adams has misunderstood the nature of the feedback processes described in our paper. We argue that under glacial conditions both the marine algae — cloud albedo and the plant-mediated CO₂ feedbacks are negative: increased temperatures cause increased albedo and increased CO₂ consumption rates, both of which tend to reduce global temperatures, damping the initial perturbation. In this regard our paper is in agreement with Adams' claim that glacial weathering rates were slower, although as we show below, some of the explanatory factors he invokes actually tend to argue for faster interglacial rates. Our feedback mechanism does tend to "pull the Earth's climate system back toward glacial level", but the "pull" is primarily from the marine algae — cloud albedo feedback loop not the land plant — CO₂ loop. Indeed, a simple numerical model (our Fig. 2) demonstrates that when the marine loop fails, the land plant loop regulates at global temperatures warmer than today's.

Many factors besides biological enhancement influence global chemical weathering rates, as Adams points out. However, in a recent analysis of the effect of changes in continental-scale hydrology and rock exposure on chemical weathering during glaciations, Gibbs and L. R. K.⁶ found that the rate of glacial chemical erosion was somewhat greater than today's, based on climate model results and a consideration of the weatherability of materials exposed on the tropical continental shelves by sea-level fall. Even if interglacial rates are higher, it is highly unlikely that they are "several times higher", as Adams contends. The depth in the ocean below which dissolution of carbonate sediments is complete (the carbonate compensation depth or CCD) is sensitive to the rate of alkalinity supply to the ocean from continental weathering⁷. The glacial sediment record of CCD variations⁸ seems to allow only small changes in chemical weathering on the timescale of glacial/interglacial fluctuations⁶. We agree that the production of highly weatherable materials (till and loess) during glaciations might produce a "burst of chemical weathering during the

interglacial conditions that follow", and may be responsible for the anomalous preservation of pelagic carbonate sediments during deglaciations (ref. 9 and J. W. Farrell, personal communication).

We apologize for not explicitly recognizing Adams' impressive work on glacial, terrestrial ecosystems in our paper. Certainly one can consider changes in the terrestrial biomass as part of the sink we describe in our abstract as the "fixation of atmospheric CO₂". We admit that our thoughts were focused on the longer-term sink associated with chemical weathering, and that on shorter (10⁴ yr) timescales changes in the terrestrial biomass have the capacity to modulate atmospheric CO₂ variations. We look forward to further stimulating discussions of the admittedly controversial ideas presented in our paper.

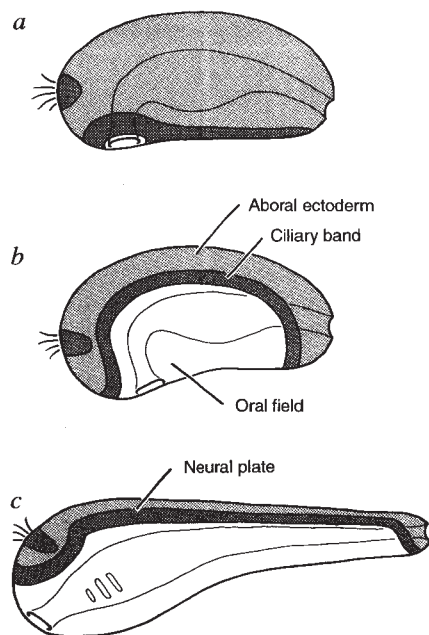
James E. Lovelock
Coombe Mill,
St Giles on the Heath,
Launceston PL 15 9RY, UK

Lee R. Kump
Earth System Science Center
and Department of Geosciences,
The Pennsylvania State
University, University Park,
Pennsylvania 16802, USA

1. Lovelock, J.E. & Kump, J.R. *Nature* **369**, 732–735 (1994).
2. Probst, J.-L. thesis, Univ. Strasbourg (1992).
3. Adams, J.M. et al. *Nature* **348**, 711–714 (1990).
4. Adams, J.M. & Woodward, F.I. *Adv. Ecol. Res.* **22**, 257–314 (1992).
5. Crowley, T.J. & North, G.R. *Paleoclimatology* (Oxford Univ. Press, 1991).
6. Gibbs, M.T. & Kump, L.R. *Paleoceanography* **9**, 529–543 (1994).
7. Delaney, M.L. & Boyle, E.A. *Paleoceanography* **3**, 137–156 (1988).
8. Farrell, J.W. & Prell, W.L. *Paleoceanography* **4**, 447–466 (1989).
9. Gibbs, M.T. & Kump, L.R. *EOS* **75**, 362 (1994).

Dorsoventral axis inversion

SIR — Arendt and Nübler-Jung¹ highlight a problem that molecular biologists have encountered when comparing the spatial expression patterns of key developmental control genes between insects and vertebrates. The axial patterns (for example, of *Hox* genes) are beautifully conserved, but patterns along the dorsoventral axis are inverted in some instances. Arendt and Nübler-Jung cite data on *decapentaplegic*, a gene involved in dorsoventral patterning, and its vertebrate homologue, *BMP-4*. Expression is dorsal in *Drosophila* and ventral in vertebrates, though there are complications: other members of this gene family in vertebrates are expressed in both the dorsal and ventral parts of the nerve cord, which is itself a dorsal structure². Another example is *wingless*, which has ventral expression in various *Drosophila* body parts³. Its vertebrate counterparts are the



Comparison of body plans in: *a*, a schematic protostome embryo (ancestral to insects); *b*, a primitive deuterostome larva (as found in echinoderms); and *c*, a chordate (vertebrate) embryo. The mouth and apical pole (shown as a ciliated apical plate) are on the left in each case, the anus is on the right. Shaded zone, area of primitive ectoderm that, in protostomes, would both cover the body and give rise to neural tissue. Darker shading, neurogenic zones, mainly ventral in protostomes. The auricularia hypothesis derives *c* from *b* by modifying the epithelial domains as shown. The neurogenic tissue of the ciliary band is incorporated into the chordate neural plate, which is then rolled up and internalized, thereby forming a dorsal nerve cord. This leaves the oral field, a tissue supposed to be new with deuterostomes, as the principal source of the tissue needed to form the definitive body surface in chordates.

Wnt genes, which are strongly expressed in the dorsal nerve cord, in most instances near the dorsal midline of that structure⁴.

The solution to this puzzle may be, as Arendt and Nübler-Jung suggest, that vertebrates have been turned upside-down during evolution, but there is an alternative and less radical way of generating at least a partial inversion. It derives from the auricularia hypothesis, a proposal for how chordates may have originated from ciliated larvae resembling those of modern echinoderms⁵. Central to the hypothesis was the idea that the chordate neural tube arose by dorsal convergence of the ciliary bands used by such larvae for locomotion (see figure). Equally important, however, is the evolutionary fate proposed for the zone of surface ectoderm known as the oral field, lying between the ciliary band and the mouth. In modern larvae of this type, it provides a surface for food capture. But in chordates, according to the auricularia hypothesis, it must presumably have taken over the whole of the external body sur-

face, because the rest of the ectoderm, dorsal to it, is supposed to be internalized by neurulation.

Very little is known for certain about the origin of the deuterostome line of animal evolution to which vertebrates and their immediate invertebrate ancestors belong. The oral field is a characteristic feature of deuterostomes, however, with no obvious counterpart in protostomes, the more ancient group to which insects belong. If the oral field is an evolutionary novelty, invented by deuterostomes, it is possible that it may have been freed to some extent from control by the fundamental systems of patterning genes active elsewhere in the ectoderm. This would provide a rationale for why gene expression patterns in insect ectoderm might map in a consistent way only to a restricted subdomain of vertebrate ectoderm, that is, to the neural plate and its derivatives, and not to the ectoderm as a whole. The map would be inverted, as required by Arendt and Nübler-Jung, because, as the neural plate rolls up, its dorsal midline is turned into ventral floorplate, while its margins, which are counterparts of the ventral-most part of the ectoderm in insects, will meet at the new dorsal midline.

The auricularia hypothesis makes no specific predictions about the patterns to be expected in surface ectoderm in vertebrates. If, however, the oral field lacks the built-in system of pattern control found elsewhere in the body, its patterning might have become more subject during evolution to secondary influences emanating from underlying mesoderm, which could then generate diverse patterns only indirectly related to those seen in surface ectoderm in other metazoa.

T. C. Lacalli

Biology Department, University of Saskatchewan,
Saskatoon, Saskatchewan,
Canada S7N-0W0

1. Arendt, D. & Nübler-Jung, K. *Nature* **371**, 26 (1994).
2. Jones, C.M. et al. *Development* **111**, 531–542 (1991).
3. Held, L.I. et al. *Wilhelm Roux Arch. dev. Biol.* **203**, 310–319 (1994).
4. Moon, R.T. *BioEssays* **15**, 91–97 (1993).
5. Garstang, W. *Zool. Anz.* **27**, 122–125 (1894).

SIR — Arendt and Nübler-Jung¹ suggest that a dorsoventral inversion occurred in the origin of the chordates, concluding that the dorsal surface of a chordate is homologous to the ventral side of non-chordates (gastreaularians), and that the dorsal nerve cord of chordates is homologous to the one or two ventral nerve cords of gastreaularians. These homologies were first proposed by Geoffroy St-Hilaire in 1822, but Arendt and Nübler-Jung now propose a mechanism by which this dorsoventral inversion could have been achieved. But a critical examination of the new evidence suggests that these purported homologies are erroneous.

Arendt and Nübler-Jung¹ argue that the ventral surface of a gastreaularian is homologous to the dorsal surface of a chordate. They support this homology with the fact that *BMP-4* is expressed ventrally in vertebrates, whereas its arthropod 'homologue' *decapentaplegic* (*dpp*) is expressed dorsally in *Drosophila*. However, *BMP-4* is not the homologue of *dpp*. A comparison of sequence similarities² suggests that the *dpp* gene duplicated generating two genes in vertebrates: *BMP-4* and *BMP-2*. What is important is that on duplication, the role of either gene may have changed and it is not clear, especially given that the expression pattern of *BMP-2* is currently unknown, which of the duplicates retained the primitive role. Hence, it is premature to conclude that the ventral side of gastreaularians and the dorsal side of chordates are homologous until the history of this gene family is known throughout the chordates.

If the dorsal surface of chordates is homologous to the ventral surface of gastreaularians, then, according to Arendt and Nübler-Jung¹, the ventral nerve cords of gastreaularians and the dorsal nerve cord of chordates are homologous. This homology must satisfy three tests³: similarity (likeness in topology, development and composition); conjunction (if feature X and feature Y are hypothesized to be homologues, but X and Y occur in the same organism, then X cannot be homologous to Y); and congruence (a parsimony test). The evidence used by Arendt and Nübler-Jung¹ is that the *achaete-scute* complex genes are expressed in neuronal precursor cells in both *Drosophila* and vertebrates. Even allowing that this evidence satisfies the test of similarity (which I would argue in part satisfies the test of similarity for nerve cells, but not nervous systems), this proposed homology does not pass the test of conjunction.

Enteropneusts, the hypothesized sister group of the chordates^{4,5}, have both a dorsal and a ventral nerve cord. Many authors have accepted the homology between the dorsal nerve cords of enteropneusts and chordates (see for example, refs 4–8), with both structures dorsal, hollow, possessing giant nerve cells^{7,9}, and formed by invagination or delamination of the neuroectoderm. The ventral nervous system of enteropneusts, is clearly of invertebrate design with circumenteric connectives and a main ventral nerve cord⁷ (suggesting that a dorsoventral inversion did not occur either in the origin of or within the deuterostomes). Because enteropneusts possess both a dorsal (X) and a ventral nerve cord (Y), the test of conjunction fails because both X and Y are found in the same organism, and therefore the proposed homology is refuted.

Two other morphological features of enteropneusts further suggest that the dorsal surfaces of enteropneusts and chordates are homologous. First, enteropneusts have a putative homologue of the chordate notochord on the dorsal side of the animal, the stomochord¹⁰. Second, enteropneusts have U-shaped branchial skeletons of identical design and orientation with chordates⁶⁻⁸.

Before one can conclude that an inversion of the dorsoventral axis occurred at the time of origin of the chordates, minimally it must be shown that: (1) *BMP-2* is expressed ventrally in vertebrates, and member(s) of the *dpp* subfamily² are expressed ventrally in 'protochordates'; (2) the dorsal nerve cord of enteropneusts and chordates are not homologues; (3) the stomochord is not the homologue of the notochord; and (4) the pharyngeal skeletons in enteropneusts and chordates arose independently. Until then, Geoffroy's original hypothesis

remains simply a matter of historical significance.

K. J. Peterson

Department of Earth and
Space Sciences and IGPP

Center for the Study of Evolution
and the Origin of Life,

University of California, Los Angeles,
California 90024-1567, USA

1. Arendt, D. & Nübler-Jung, K. *Nature* **371**, 26 (1994).
2. Kingsley, D. M. *Genes Dev.* **8**, 133-146 (1994).
3. Patterson, C. in *Problems of Phylogenetic Reconstruction* (eds Joysey, K. A. & Friday, A. E.) 21-74 (Academic, London, 1982).
4. Peterson, K. J. in *Major Features of Vertebrate Evolution* (eds Prothero, D. R. & Schoch, R. M.) 14-37 (Paleont. Soc. Univ. Tennessee, Knoxville, 1994).
5. Schram, F. R. in *The Early Evolution of Metazoa and the Significance of Problematic Taxa* (eds Simonetta, A. M. & Conway Morris, S.) 35-46 (Cambridge Univ. Press, 1991).
6. Brusca, R. C. & Brusca, G. J. *Invertebrates* (Sinauer, Sunderland, Massachusetts, 1990).
7. Hyman, L. H. *The Invertebrates* Vol. 5 (McGraw-Hill, New York, 1959).
8. Nielsen, C. *Biol. J. Linn. Soc.* **25**, 243-299 (1985).
9. Sarnat, H. B. & Netsky, M. G. *Evolution of the Nervous System*, 2nd edn (Oxford Univ. Press, 1981).
10. Balser, E. J. & Ruppert, E. E. *Acta zool., Stockh.* **71**, 235-249 (1990).

Cation selectivity in ion channels

SIR — Valera *et al.*¹ and Brake *et al.*² have reported the amino-acid sequences of members of a novel class of ligand-gated ion channel, namely the ATP-gated channel or P_{2x} receptor. The channel dis-

in alignment of the highly conserved TTTXGXG motif in the H5 region of potassium channels⁵. Of particular interest is the

inwardly-rectifying and voltage-gated potassium channels tyrosine is conserved in the central position. In the *ether-à-go-go* channel⁶, tyrosine is replaced by phenylalanine, while in the P_{2x} receptor^{1,2} serine or valine present.

As proposed by Kumpf and Dougherty⁷, the presence of aromatic amino acids within H5 may result in potassium-selective 'cation- π ' interactions in potassium channels. Thus, highly potassium-selective inward-rectifiers and voltage-gated channels exhibit a conserved tyrosine in GXG (or phenylalanine in *ether-à-go-go*), whereas in the non-specific cation channels encoded by P_{2x} tyrosine is replaced by a smaller, non-aromatic residue^{1,2}.

I.D. Kerr

M.S.P. Sansom

Laboratory of Molecular Biophysics,
Rex Richards Building,
University of Oxford,
South Parks Road,
Oxford OX1 3QU, UK

1. Valera S. *et al.* *Nature* **371**, 516-519 (1994).
2. Brake, A. J. *et al.* *Nature* **371**, 519-524 (1994).
3. Kubo, Y. *et al.* *Nature* **362**, 127-133 (1993).
4. Barton, G. J. & Sternberg, M. J. E. *J. molec. Biol.* **198**, 327-337 (1987).
5. Heglinotham, L. *et al.* *Biophys. J.* **66**, 1061-1067 (1994).
6. Brüggenmann, A. *et al.* *Nature* **365**, 445-448 (1993).
7. Kumpf, R. A., & Dougherty, D. A. *Science* **261**, 1708-1710 (1993).

Bet on positional information

SIR — One idea of how patterns of cellular differentiation are specified during development is that first, positional information is set up in a population of cells, and second, that each cell interprets this information according to its genetic constitution and developmental history¹. Recent studies have suggested that a related set of homeotic genes encode positional identity along the body axis of many animals, and it has even been suggested that this defines a zootype that all animals share². In insects, for example, one can think of the homeotic genes as acting together to give genetic addresses, which establish segment identity^{3,4}. Martinez Arias⁵ has challenged these ideas on the basis of the work of Warren *et al.*, who have studied the expression of the homeotic gene *Ultrabithorax* (*Ubx*) in butterflies⁶.

In *Drosophila*, *Ubx* is expressed in the third thoracic segment which forms a haltere. In butterflies the third thoracic segment carries a wing. According to Martinez Arias, "most punters would have put their money on the absence of expression" of *Ubx* in the third thoracic segment of butterflies because he thinks the homeotic genes are associated with specific structures. But I, and other colleagues I have asked, would not have put down even a penny in support of such an idea, for the identity of a segment as specified by homeotic genes says nothing about what structures it will develop. That is the central idea of positional information. Indeed Warren *et al.*'s finding that in butterflies the third thoracic segment expresses *Ubx* strongly supports the idea of positional identity being specified by the homeotic genes. What has changed in evolution of the basic pattern of butterflies and flies is not the positional identity of the thoracic segments nor the deployment of homeotic genes, but their downstream targets.

Lewis Wolpert

Department of Anatomy &
Developmental Biology,

University College London Medical School,
Cleveland Street, London W1P 6DB, UK

1. Wolpert, L. *Development* (Suppl.) 3-12 (1989).
2. Slack, J. M. W. *et al.* *Nature* **361**, 490-492 (1993).
3. Lawrence, P. A. *The Making of a Fly* (Blackwell Scientific, Oxford, 1992).
4. Lawrence, P. A. & Morata, G. *Cell* **35**, 595-601 (1983).
5. Martinez Arias, A. *Nature* **372**, 408-409 (1994).
6. Warren, R. W. *et al.* *Nature* **372**, 458-461 (1994).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references.

	1	2	3	4	5	6	7	8	9	10	*							
IRK1 K ⁺ channel	T	A	A	F	L	F	S	I	E	T	T	T	T	G	Y	G	F	R
IRK2 K ⁺ channel	M	A	A	F	L	F	S	I	E	T	T	T	T	G	Y	G	L	R
KATP K ⁺ channel	V	S	A	F	L	F	S	I	E	T	T	T	T	G	Y	G	F	R
ROMK1 K ⁺ channel	T	S	A	F	L	F	S	I	E	T	T	T	T	G	Y	G	F	R
Shaker A K ⁺ channel	P	D	A	F	W	W	A	V	T	M	T	T	V	G	G	D	M	
DRK1 K ⁺ channel	P	A	S	F	W	W	A	T	I	T	M	T	T	V	G	G	D	I
<i>ether-a-go-go</i>	V	T	A	L	Y	F	T	M	T	C	M	T	S	V	G	F	G	N
<i>P_{2x}</i> PC12 cells	I	P	T	I	I	N	L	A	T	A	L	T	S	I	G	V	G	S
<i>P_{2x}</i> vas deferens	K	A	G	K	F	D	I	I	P	T	M	T	T	G	S	G	I	G

Alignment of the H5 region of inwardly-rectifying potassium channels (IRK1, IRK2, KATP and ROMK1), voltage-gated potassium channels (DRK1, Shaker A and *ether-à-go-go*) and ATP-gated channels (P_{2x} receptors). Boxed region, highly conserved eight-residue potassium-channel signature sequence⁵. Asterisk, denotes the central position of the GXG motif.

plays cation selectivity when expressed in *Xenopus* oocytes, and both groups of authors propose a transmembrane topology similar to that of a recently cloned family of inwardly-rectifying potassium channels³, with two potential membrane-spanning helices (M1 and M2) and an intervening region, H5. The sequence alignment proposed by Valera *et al.*¹ for P_{2x} cloned from rat vas deferens is very similar to inward-rectifiers and voltage-gated potassium channels within its H5 region.

We would like to propose an alternative assignment for the H5 region of P_{2x} cloned from rat PC12 cells to that offered by Brake *et al.*². Our assignment was obtained by alignment of the proposed transmembrane regions of voltage-gated and inwardly-rectifying potassium channels with the P_{2x} sequences^{1,2}, using the multiple sequence alignment package, AMPS⁴. Our analysis (see figure) results

Samurai of the cosmos

William H. Press

Voyage to the Great Attractor: Exploring Intergalactic Space. By Alan Dressler. Knopf: 1994. Pp. 355. \$25.

THE observatories of the Carnegie Institution, based in Pasadena, California, are home not only to Alan Dressler, whose elegant musings about the nature of science and discovery are the heart of this book, but also to (among others) Allan Sandage, whose recorded lucubrations in Dennis Overbye's 1991 *Lonely Hearts of the Cosmos* mark perhaps the greatest unintentional *auto da fé* in the entire scientific literature, and Wendy Freedman, a brilliant younger observer at the centre of new cosmological discoveries made with the Hubble Space Telescope. If the Universe has a centre, it may well turn out to be on Santa Barbara Street.

Dressler's book, however, seems more informed by the novels of Jules Verne and the plans hatched in Phileas Fogg's Reform Club than by any lazy California sunshine. He promises the reader a rollicking good story about a motley band of astronomical explorers (others nicknamed them the 'Seven Samurai'), and, for the most part, he delivers on his promise. Indeed, a lot of the charm of this book is in the portraits of the seven (a group diverse in nationality, gender, age and temperament) and the conflicts, internal and external, engendered by their scientific collaboration.

The decade from the early 1980s to the early 1990s was an extraordinarily fertile time for cosmology. Although his native expertise lies in observation, Dressler's explanations of both theoretical and observational matters are equally lucid and down to earth. He develops a good set of analogies to everyday experience, and uses them consistently throughout his book.

On the theoretical side, the decade saw the development of Alan Guth's conceptualization of inflationary cosmology; its consequences, that the Universe should be almost closed ($\Omega=1$ in technical shorthand) and that it must have a certain relative balance ('spectrum') of large-scale and small-scale matter perturbations; the improved evidence, from ratios of isotopes produced in the Big Bang, that there are too few protons, neutrons or electrons to make up the mass required by inflation — hence the need to posit the idea of 'dark matter'; the development of a detailed theory for the formation of structure — galaxies and clusters of galaxies — in the presence of cold dark matter, the so-called 'CDM model'; and the prediction of the tiny perturbations that the seeds of the CDM model should make on

the otherwise homogeneous cosmic microwave background.

The observational side repeatedly demonstrated that, whatever the excesses of theorists, nature was even more bizarre. Huge voids, virtually empty of galaxies, were found in space. Initially these were thought to be rare anomalies. But then Margaret Geller, John Huchra and Valerie de Lapparent showed that voids were the rule not the exception: most galaxies are concentrated in relatively thin 'walls' between the voids. This bubble-like pattern is the Universe's 'large-scale structure'. How could earlier observers have missed this? The answer is that the third dimension — depth — is extremely difficult for astronomers to measure at extragalactic distances. With modest errors in depth and (initially) a paucity of data, the voids simply did not stand out.

Initially distinct from the discovery of large-scale structure and started before the scope of that structure was known, it is the work of the Seven Samurai (Sandra Faber, Donald Lynden-Bell, Dressler, Roberto Terlevich, Daniel Burstein, Gary Wegner and Roger Davies) that furnishes the main plot of the Jules-Vernean adventure story within Dressler's book. This group's project began as an ambitious but prosaic effort systematically to survey large numbers of so-called giant elliptical galaxies. Serendipitously, they discovered a subtle way of measuring the distance to their individual galaxies — depth, again — independently of the galaxies' recession velocities (redshifts). The redshift of a galaxy can then be used as a measure of its 'peculiar velocity' or deviation from a smooth expansion velocity. To their astonishment, and to the initial disbelief of the astronomical world, the Seven Samurai found a coherent pattern of 'large-scale flows', now generally appreciated (although in a way not yet fully understood) to be the coherent, gravity-driven velocities responsible for creating the bubble-like large-scale structure in the first place.

In 1991, a COBE (Cosmic Background



GOLD retriever — an alchemist with a flask and scroll that reads *Eamus Quæsitum Quasuos Elementorum [sic]* ("let us go in search of the natures of the four elements"). From Solomon Trismosin's *Splendor Solis* (1582). The picture appears in *The Mirrors of Alchemy* by Gareth Roberts, a beautifully presented account of alchemical ideas and images from antiquity to the seventeenth century. The book draws heavily on the outstanding collection of manuscripts and books in the British Library. Published by the British Library, £22.50.

Explorer) satellite team led by George Smoot crowned these ground-based achievements by detecting minute imprints on the cosmic microwave background, at a level and on a scale generally consistent with both the CDM prediction and the reality of the large-scale structure and the large-scale flows.

I say "generally consistent" because the past three years have seen considerable wrangling over whether the diverse parts of the cosmology puzzle truly fit together or whether additional work with a theoretical hacksaw and file is necessary. Dressler takes the moderate, and I think correct, view that the story so far is remarkable for its success, even if it is not yet complete. Indeed, right now, both theorists and observers are filing away at a poor fit between the value of the Hubble constant, the apparent age of globular cluster stars and the continuing belief (and some evidence from the large-scale flows) that $\Omega = 1$.

Although he may never again have a story to tell with a plot as compelling as his own adventures with the six other Samurai, one hopes that Dressler will continue in further charming books to recount the saga of modern cosmological discovery as it continues to unfold. □

William H. Press is at the Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA.

Drag artist

Rory Howlett

Life in Moving Fluids: The Physical Biology of Flow, Second Edition. By Steven Vogel. Princeton University Press: 1994. Pp. 467. \$49.50, £35.

FORM and function are the meat and drink of comparative biology, but a whole generation of factory-trained molecular technicians are poorly equipped to appreciate the morphological subtleties of the animals and plants they purport to study. Ask the average *Drosophila* developmental geneticist why his or her beloved fruit-flies devote just half of one per cent of their body mass to wings and you are unlikely to be any the wiser after hearing the answer. Those who wish to know the answer to this and a vast array of related questions should turn to this fascinating book — now in its second edition — in which Steven Vogel, a biologist in the true sense of the word, explores the sometimes turbulent interface between fluid dynamics and biology, with examples ranging from bacteria to brachiopods, from kelp to killer bees.

The subject of the book forms a branch of biophysics in which the polymerase chain reaction is about as useful as a patent office would be to a human geneticist with a belief in the sanctity of the

genome. It is a field where researchers build ingenious devices for the measurement of tiny pressure differentials in inaccessible places, minimum and maximum flow velocities, drag coefficients and the like. Moreover, it is fun: stuffed birds sprayed with hairspray are mounted in wind tunnels to study the effect of feather fluttering on drag and lift; scale models of prairie dog burrows are built to ventilate questions of passive airflow; penguin carcasses are dragged at various speeds behind motor launches to measure drag, with and without wings; a glass tube with an enlargement just upstream of the final orifice is used to illustrate how both the human penis and ink-jet printers are designed to minimize the formation of those inconvenient 'satellite droplets'.

The last example highlights the analogies that can often be drawn between the functional morphology of organisms and the design of human-built contraptions: engineering considerations and physical concepts are to the fore — and it is here that the typically innumerate biologist waves his or her arms in horror and retreats to more qualitative territories. Much of the value of Vogel's book is that it provides a clear and accessible account of physical measurement, theory and analysis before diving into the detailed intricacies of specific biological examples. Starting with the relatively simple examples that assume laminar, nonturbulent flow, he guides the reader through the basic principles of fluid dynamics, including streamlines, continuity of flow, the role of pressure and momentum, surface effects and velocity gradients.

Throughout, attention is drawn to the sometimes paradoxical behaviour of fluids and to the importance of the many dimensionless ratios, such as the ubiquitous Reynolds number, describing the relative importance of inertial and viscous forces in any given biological situation, whether it be the flow of blood through the vascular system of vertebrates or the dispersal of fungal spores. More complex phenomena such as turbulence and vortex shedding are handled deftly, leaving the reader confident that he or she need not despair even in the face of biological problems characterized by complex, nonsteady flows.

All in all, this is a fine book, with an extensive up-to-date reference list, and written with enthusiasm and humour. It is heartily recommended to all serious students of biology, and especially to those who sympathize with Vogel's sentiment, aired as a coda, that although we may live in an age of great advance in the biological sciences, progress per investigator may be at a low point, with faddishness to blame. □

Rory Howlett is an associate editor of Nature.



A DESERT millipede *Orthoporus ornatus* on scrubland near the Rio Bravo in Coahuila, Mexico. This picture, and the one on the facing page, are taken from *Two Eagles/Dos Aguilas*, a magnificent photographic celebration of the natural world of the United States/Mexico borderlands. Published by University of California Press in association with the Nature Conservancy, \$55.

Adjudicating science

Lee Loevinger

Reference Manual on Scientific Evidence. By William W. Schwarzer *et al.* Federal Judicial Center, Dolley Madison House, 1520 H Street NW, Washington DC 20005, USA: 1994. Pp. 637. \$25 (pbk).

THE Federal Judicial Center is an agency for policy research and education within the judicial branch of the US government. In 1990, in response to the increasingly complex and esoteric issues coming before federal courts, it decided to prepare a manual to assist judges in managing expert testimony, primarily on issues of science and technology. This volume is the result. As a government publication, the manual is not subject to copyright, and will be made available this month to the principal lawbook publishers, including computer legal networks.

The manual is divided into three parts, the first and third of which each consist of two articles on legal aspects of scientific testimony. The first article discusses the procedural devices by which courts may define and refine technical issues and manage the presentation of expert testimony under the Federal Rules of Civil Procedure. The second article discusses the conditions for admissibility of scientific testimony, which involve both the qualifications of witnesses and the foundation and nature of the testimony. With regard to the latter issue, the decision made in June 1993 by the Supreme Court in the case of *Daubert v. Merrell Dow Pharmaceuticals* established an entirely new set of rules for federal courts, the application of which is still being developed. The discussion of the standards of scientific validity now required for the admissibility of scientific evidence relies largely on cases decided before that date, and so does not add much to the general principles enunciated in the Supreme Court opinion.

The second and principal part of the manual contains seven 'reference guides' detailing the fundamental scope and nature of specific areas of science. The subjects covered are epidemiology, toxicology, survey research, forensic DNA evidence, statistics, multiple regression and the estimation of economic losses in damages awards. Because each article is written by a different author, or set of authors, there is some variation in both style and approach. But the differences are not as great as might be expected in an anthology: each guide has a common framework, with an outline preceding the text, which in turn is divided into sections on issues or subjects peculiarly relevant to



MEMBERS of a 1907 expedition to the Pinacate desert region of northern Sonora, Mexico, extract water from a cactus. From *Two Eagles/Dos Aguilas*, with photographs by Tupper Ansel Blake and text by Peter Steinhart. For details see page 114.

the field; substantial cross-referencing; jargon expounded in reasonably comprehensible language; a glossary that defines the technical patois of the field; and a bibliography of additional reference sources.

The seven guides do not, and could not, attempt to make experts out of the reader; but a diligent reader will be educated sufficiently to converse intelligently with an expert in a particular field and to ask questions that would determine the validity of the expert's testimony. Indeed, much of each guide consists of an explanation of the limitations of expertise in the field, and all are strictly expository rather than anecdotal or philosophical.

As might be expected, a common theme is statistics, which are used in most areas of empirical science. Correspondingly, the longest guide is given over to this subject, and is supplemented by a comprehensive glossary of terms and by another article that illustrates and explains in mathematical detail the statistical process of multiple regression. These sections should be useful to anyone interested in any field of science or technology.

The third part of the manual consists of two articles on special procedures that may aid the consideration of scientific evidence by a federal court. One reports the results of a study by the Federal Judicial Center of court-appointed experts acting under rule 706 of the Federal Rules of Evidence. This permits a federal judge to appoint a presumably impartial expert in a case appearing to require such a procedure. This authority has been used infrequently but successfully in a few cases.

The last article describes a similar study of the power of federal judges to appoint a special master (or referee) under Federal Rule of Civil Procedure 53 to report on a particular aspect of a complex case. This power, too, is rarely used, although it is potentially useful.

The introductory and concluding articles are specifically oriented toward trials in federal courts under existing rules, and will be of interest primarily to federal judges and trial lawyers. These do not have accompanying glossaries. But the seven substantial guides that deal with areas of science relevant to contemporary legal problems will be of interest to anyone who seeks information about the subjects covered. They do not make for particularly easy reading, but all of them are illuminating; the text and the glossary will help in understanding not only testimony but also other technical treatises in the field.

The manual does not cover all types of scientific data that may be used in legal proceedings. Many subjects, such as physics, chemistry and psychiatry, are not included; but future, and no doubt expanded, editions are planned. The scientific expositions provide more questions than answers. In both science and law, however, the first step toward knowledge is to ask the right questions; and in this respect the manual will provide invaluable assistance. □

Lee Loevinger is at Hogan and Hartson, 555 Thirteenth Street NW, Washington DC 20004, USA.

Mam and Sam

Douglas Palmer

Mammoth. By Adrian Lister and Paul Bahn. Macmillan Inc., USA: 1994. Pp. 168. \$30.

Seismosaurus: The Earth Shaker. By David G. Gillette. Columbia University Press: 1994. Pp. 205. \$39.95, £26.50.

'MAMMOTH' is still a remarkably and perhaps peculiarly common word that continues to be used to epitomize immensity despite our general knowledge of much larger animals (blue whales, sauropod dinosaurs such as *Seismosaurus*).

But it is perhaps no accident that we have this linguistic association, given our past material and imaginative connection with mammoths and the other members of the Ice Age megafauna. As this book so clearly illustrates and describes, for more than 200,000 years these shaggy elephants were a dominant species in the Northern Hemisphere, with herds constantly on the move, grazing their way back and forth with the seasons. And, like the vast herds of buffalo that roamed the Great Plains of North America, they seem to have been too much of a temptation for human hunters to resist.

From boomerangs (see *Nature* 329, 436; 1987) to building blocks, clothes to fuel, the bits and pieces of mammoths have been put to use by human hunters, their form evoked on cave walls and in sculpture; but whether or not our ancestors finally wiped them out is still a moot point. As has been recently discovered (*Nature* 362, 337; 1993), mammoths survived on the Arctic island of Wrangel until as recently as 4,000 years ago, becoming dwarfed in the process. As the authors show, their extinction is a complex issue. Read the book and decide for yourself.

Like the best of semipopular books, this is an excellent example of the soft approach to serious science: it is accurate and up to date (the Wrangel and other 1993 information is included) and written by acknowledged international experts in the field. Buy it to encourage a youngster but read it yourself first — mammoths can

Long views of the past — from left to right, *Seismosaurus*, *Ultrasaurus* and *Supersaurus*. Monochrome painting by John Sibbick, 1992.

still engage an adult's imagination.

The most frustrating thing about these remarkable beasts is that we have missed them by only some 4,000 years. There is a lesson here for our treatment of the remaining megafauna: it would seem that apparently isolated populations of a few thousand are not enough for the survival of a species.

We have had *Supersaurus* and *Ultrasaurus*, but even in the media-driven world of dinosaur superlatives, *Seismosaurus* (Earth-shaker lizard) must be near the apogee in general largesse at between 128 and 170 feet in length and about 100 tonnes in weight.

Fortunately, Gillette avoids the temptation to indulge in the near-hysteria so often fostered by 'dino hype' and tells an interesting and straightforward tale of his excavation of this remarkable sauropod and its taphonomy (study of burial processes).

Originally found in 1979 by some hikers in the mesalands of New Mexico near Albuquerque, *Seismosaurus* is by far the longest dinosaur known. The story has taken the intervening 15 years to mature. Considering the potential of such an interesting find to generate rancour and skulduggery between vested interests, it is heartening to find the unfolding story to be one of remarkable luck, persistence, cooperation and painstaking work by a lot of people, both professionals and amateurs in the true sense.

Luckily the finders included a geologist, Jan Cummings, who had a good idea of what had been found, a few unusually large and still connected vertebrae of a sauropod dinosaur in the late Jurassic sandstones of the Morrison formation. He was able to convince Gillette, then curator of the New Mexico Museum of Natural History in Albuquerque, to have a look. Luckily again, Gillette was able to engage the interest and expertise of a whole range

of scientists from the Los Alamos, Oak Ridge and Sandia National Laboratories. They were able to bring enormous enthusiasm for problem-solving and to gain access to high-tech equipment for finding more of 'Sam', the seismic sauropod.

Hydroxyapatite, the mineral component of bone, fluoresces, so they tried hunting at night with ultraviolet light; fossil bone seems to concentrate uranium minerals, so they tried high-resolution gamma-ray detectors; then there was ground-penetrating radar, proton magnetometry and acoustic-diffraction tomography, all of which were partially successful in predicting where more bone material might be buried. But it still required traditional excavation techniques to recover what remained in the rock. All these techniques are simply and effectively explained for the novice and the book is an excellent and salutary stimulus for dinosaur enthusiasts of any age.

Finally, the original population of this dinosaur must have been in the order of thousands, with the species surviving for a million years or so. The total number of potential fossil skeletons, each with several very large and heavy bones, must therefore be in the order of millions. And yet we have only one partial skeleton, without the large limb bones. Does this tell us something important about the vagaries of the fossilization process and loss of information about past life? Just how good is the sample? There is nothing novel about these questions but they do emphasize the importance of such remarkable and apparently 'one-off' finds well beyond their statistical insignificance. That these discoveries can be made by anyone with a 'good eye' and a bit of luck gives them a lure that helps to keep palaeontology going. □

Douglas Palmer is at 31 Mawson Road, Cambridge CB1 2DZ, UK.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Natural History Museum/M. Long

Viral dynamics in human immunodeficiency virus type 1 infection

Xiping Wei^{*}, Sajal K. Ghosh^{*}, Maria E. Taylor^{*}, Victoria A. Johnson[†], Emilio A. Emini[‡], Paul Deutsch[§], Jeffrey D. Lifson^{||}, Sebastian Bonhoeffer[¶], Martin A. Nowak^{||}, Beatrice H. Hahn^{*}, Michael S. Saag[†] & George M. Shaw^{*#}

Divisions of ^{*} Hematology/Oncology and [†] Infectious Diseases, University of Alabama at Birmingham, 613 Lyons-Harrison Research Building, 701 South 19th Street, Birmingham, Alabama 35294, USA

Departments of [‡] Antiviral Research and [§] Clinical Pharmacology, Merck Research Laboratories, West Point, Pennsylvania 19486, USA

^{||} Division of HIV and Exploratory Research, Genelabs Technologies Inc., Redwood City, California 94063, USA

[¶] Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

The dynamics of HIV-1 replication *in vivo* are largely unknown yet they are critical to our understanding of disease pathogenesis. Experimental drugs that are potent inhibitors of viral replication can be used to show that the composite lifespan of plasma virus and virus-producing cells is remarkably short (half-life ~2 days). Almost complete replacement of wild-type virus in plasma by drug-resistant variants occurs after fourteen days, indicating that HIV-1 viraemia is sustained primarily by a dynamic process involving continuous rounds of *de novo* virus infection and replication and rapid cell turnover.

THE natural history and pathogenesis of human immunodeficiency virus type-1 (HIV-1) infection are linked closely to the replication of virus *in vivo*¹⁻¹⁷. Clinical stage is significantly associated with all measures of virus load, including infectious virus titres in blood, viral antigen levels in serum, and viral nucleic acid content of lymphoreticular tissues, peripheral blood mononuclear cells (PBMCs) and plasma (reviewed in ref. 18). Moreover, HIV-1 replication occurs preferentially and continuously in lymphoreticular tissues (lymph node, spleen, gut-associated lymphoid cells, and macrophages)^{11,19,20}; virus is detectable in the plasma of virtually all patients regardless of clinical stage^{6,10,13,21}; and changes in plasma viral RNA levels predict the clinical benefit of antiretroviral therapy (R. Coombs, unpublished results). These findings emphasize the central role of viral replication in disease pathogenesis.

Despite the obvious importance of viral replication in HIV-1 disease, relatively little quantitative information is available regarding the kinetics of virus production and clearance *in vivo*, the rapidity of virus and CD4⁺ cell population turnover, and the fixation rates of biologically relevant viral mutations^{22,23}. This circumstance is largely due to the fact that previously available antiretroviral agents lacked sufficient potency to abrogate HIV-1 replication, and methods to quantify virus and determine its genetic complexity were not sufficiently sensitive or accurate. We overcame these obstacles by treating subjects with new investigational agents which potently inhibit the HIV-1 reverse transcriptase (nevirapine, NVP)²⁴ and protease (ABT-538; L-735,524)^{25,26}; by measuring viral load changes using sensitive new quantitative assays for plasma virus RNA^{6,18,27}; and by quantifying changes in viral genotype and phenotype in uncultured plasma and PBMCs using automated DNA sequencing²⁸ and an *in situ* assay of RT function^{29,30}.

Virus production and clearance

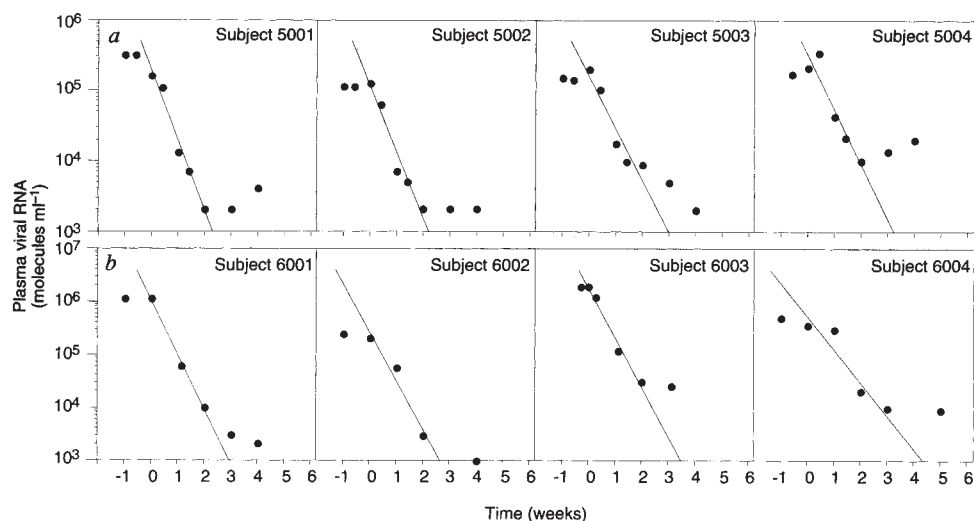
Twenty-two HIV-1-infected subjects with CD4⁺ lymphocyte counts between 18 and 251 per mm³ (mean $\pm$ s.d., 102 ± 75 cells per mm³) were treated with ABT-538 ($n=10$), L-735,524 ($n=8$) or NVP ($n=4$) as part of phase I/IIA clinical studies. The

design and clinical findings of those trials will be reported elsewhere (K. Squires *et al.*, and V.A.J. *et al.*, manuscripts in preparation). Plasma viral RNA levels in the 22 subjects at baseline ranged from $10^{4.6}$ to $10^{7.2}$ molecules per ml (geometric mean of $10^{5.5}$) and exhibited maximum declines generally within 2 to 4 weeks of initiating drug therapy (Figs 1 and 2a). For ABT-538- and L-735,524-treated patients, virus titres fell by as much as $10^{3.9}$ -fold (mean decrease of $10^{1.9}$ -fold) whereas for NVP-treated patients virus fell by as much as $10^{2.0}$ -fold (mean decrease of $10^{1.6}$ -fold). The overall kinetics of virus decline during the initial weeks of therapy with all three agents corresponded to an exponential decay process (Figs 1 and 2a).

The antiretroviral agents used in this study, despite their differing mechanisms of action, have a similar overall biological effect in that they block *de novo* infection of cells. Thus the rate of elimination of plasma virus that we measured following the initiation of therapy is actually determined by two factors: the clearance rate of plasma virus *per se* and the elimination (or suppression) rate of pre-existing, virus-producing cells. To a good approximation, we can assume that virus-producing cells decline exponentially according to $y(t) = y(0)e^{-\alpha t}$, where $y(t)$ denotes the concentration of virus-producing cells at time t after the initiation of treatment and α is the rate constant for the exponential decline. Similarly, we assume that free virus $v(t)$ is generated by virus-producing cells at the rate $ky(t)$ and declines exponentially with rate constant u . Thus, for the overall decline of free virus, we obtain $v(t) = v(0)[ue^{-\alpha t} - \alpha e^{-u t}]/(u - \alpha)$. The kinetics are largely determined by the slower of the two decay processes. As we have data only for the decline of free virus, and not for virus-producing cells, we cannot determine which of the two decay processes is rate-limiting. However, the half-life ($t_{1/2}$) of neither process can exceed that of the two combined. With these considerations in mind, we estimated the elimination rate of plasma virus and of virus-producing cells by three different methods: (1) first-order kinetic analysis of that segment of the viral elimination curve corresponding to the most rapid decline in plasma virus, generally somewhere between days 3 and 14; (2) fitting of a simple exponential decay curve to all viral RNA determinations between day 0 and the nadir or inflection point (Fig. 1); and (3) fitting of a compound decay curve

[#] To whom correspondence should be addressed.

FIG. 1 Plasma viral RNA determinations in representative subjects treated with the HIV-1 protease inhibitors ABT-538 (a) and L-735,524 (b). Subjects had not received other antiretroviral agents for at least 4 weeks before therapy. Treatment was initiated at week 0 with 400–1,200 mg d⁻¹ of ABT-538 or 1,600–2,400 mg d⁻¹ of L-735,524 and was continued throughout the study. Viral RNA was determined by modified branched DNA (bDNA)¹⁸ (a) or RT-PCR²⁷ (b) assay and confirmed by QC-PCR⁶. Shown are the least-squares fit linear-regression curves for data points between days 0 and 14 indicating exponential (first-order) viral elimination.



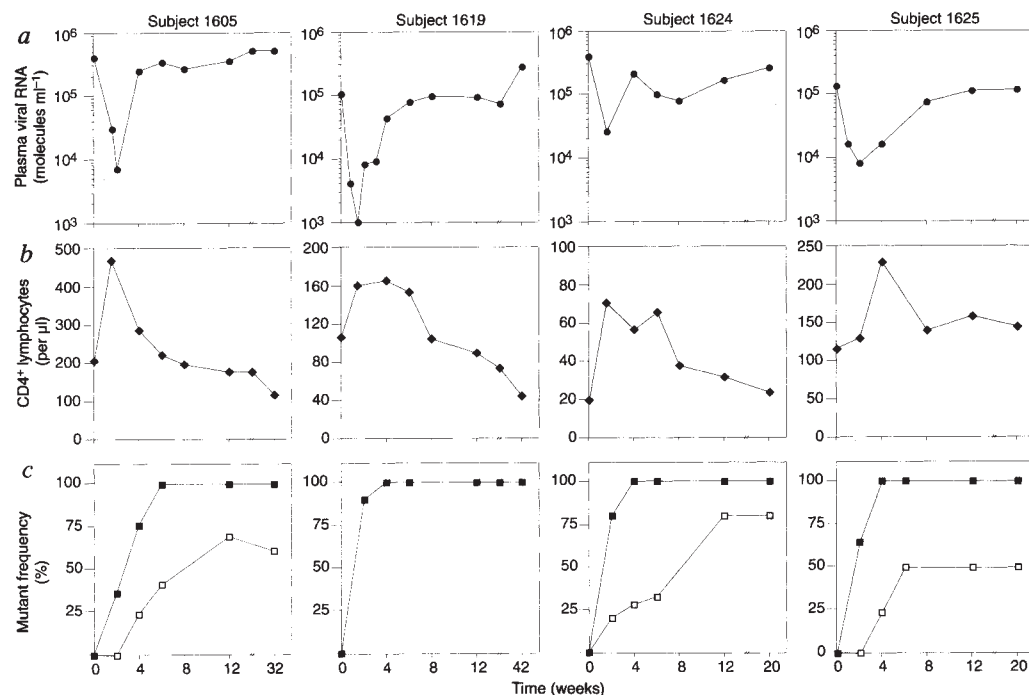
that takes into account the two separate processes of elimination of free virus and virus-producing cells, as described. Method (1) gives a $t_{1/2}$ of 1.8 ± 0.9 days; method (2) gives a $t_{1/2}$ of 3.0 ± 1.7 days; and method (3) gives a $t_{1/2}$ of 2.0 ± 0.9 days for the slower of the two decay processes and a very similar value, 1.5 ± 0.5 days, for the faster one. These are averages (± 1 s.d.) for all 22 patients. Method (3) arguably provides the most complete assessment of the data, whereas method (2) provides a simpler interpretation (but slightly slower estimate) for virus decline because it fails to distinguish the initial delay in onset of antiviral activity due to the drug accumulation phase, and the time required for very recently infected cells to initiate virus expression, from the subsequent phase of exponential virus decline. There were no significant differences in the viral clearance rates in subjects treated with ABT-538, L-735,524 or NVP, and there was also no correlation between the rate of virus clearance from

plasma and either baseline CD4⁺ lymphocyte count or baseline viral RNA level.

Virus turnover

Direct population sequencing. As an independent approach for determining virus turnover and clearance of infected cells, we quantified serial changes in viral genotype and phenotype with respect to drug resistance in the plasma and PBMCs of four subjects treated with NVP (Fig. 2). NVP potently inhibits HIV-1 replication but selects for one or more codon substitutions in the reverse transcriptase (RT) gene^{24,31,32}. These mutations result in dramatic decreases (up to 1,000-fold) in drug susceptibility and are associated with a corresponding loss of viral suppression *in vivo*³². Genetic changes resulting in NVP resistance can thus serve as a quantifiable molecular marker of virus turnover. A rapid decline in plasma viral RNA was

FIG. 2 Plasma viral RNA determinations (a), CD4⁺ lymphocyte counts (b), and percentages of mutant viral genomes in plasma and PBMCs (c) of subjects initiating treatment with NVP. Subjects were participants in a clinical protocol assessing the effects of NVP when added to existing treatment with ddI (subject 1605) or ddI plus zidovudine (subjects 1619, 1624, 1625). Treatment with NVP was initiated at week 0 using 200 mg per day and was increased to 400 mg per day after 2 weeks. ddI and zidovudine dosages were 400 mg per day and 300–600 mg per day, respectively. Viral RNA (●) was determined by QC-PCR assay⁶. CD4⁺ lymphocytes (◆) were quantified by flow cytometry. Frequencies of viral genomes containing NVP-resistance-associated mutations in plasma (■) and PBMCs (□) were determined by automated DNA sequence analysis (Fig. 3, legend), with each data point representing the average of 3–6 independent PCR amplifications and sequence determinations.



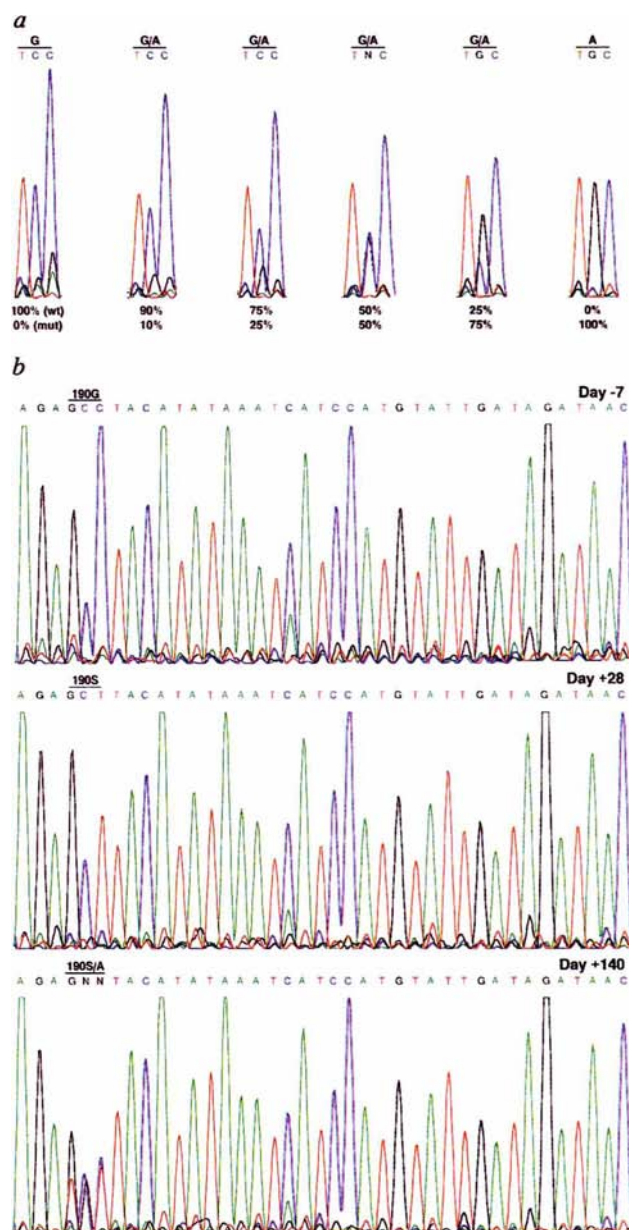
observed following the institution of NVP therapy and this was associated with a reciprocal increase in CD4⁺ lymphocyte counts (Fig. 2a and b). Both responses were of limited duration, returning to baseline within 6–20 weeks in these four patients. The proportion of virus in uncultured plasma and PBMCs that contain NVP-resistance-conferring mutations (Fig. 2c) was determined by direct automated nucleotide sequencing of viral nucleic acid (Fig. 3), as previously described²⁸. We first validated this method by reconstitution experiments, confirming its sensitivity for detecting RT mutants that comprise as little as 10% of the overall virus population. Defined mixtures of wild-type and mutant HIV-1 RT cDNA clones (differing only at the second base position of codon 190) were amplified and sequenced (Fig. 3a). Varying proportions of wild-type and mutant viral sequences present in the original DNA mixtures (mutant composition: 0, 10, 25, 50, 75 and 100%) were faithfully represented in the relative peak-on-peak heights (and in the relative peak-on-peak areas) of cytosine (C) and guanine (G) residues at the second base position within this codon. Ratios of (mutant)/(mutant + wild type) nucleotide peak heights expressed in arbitrary

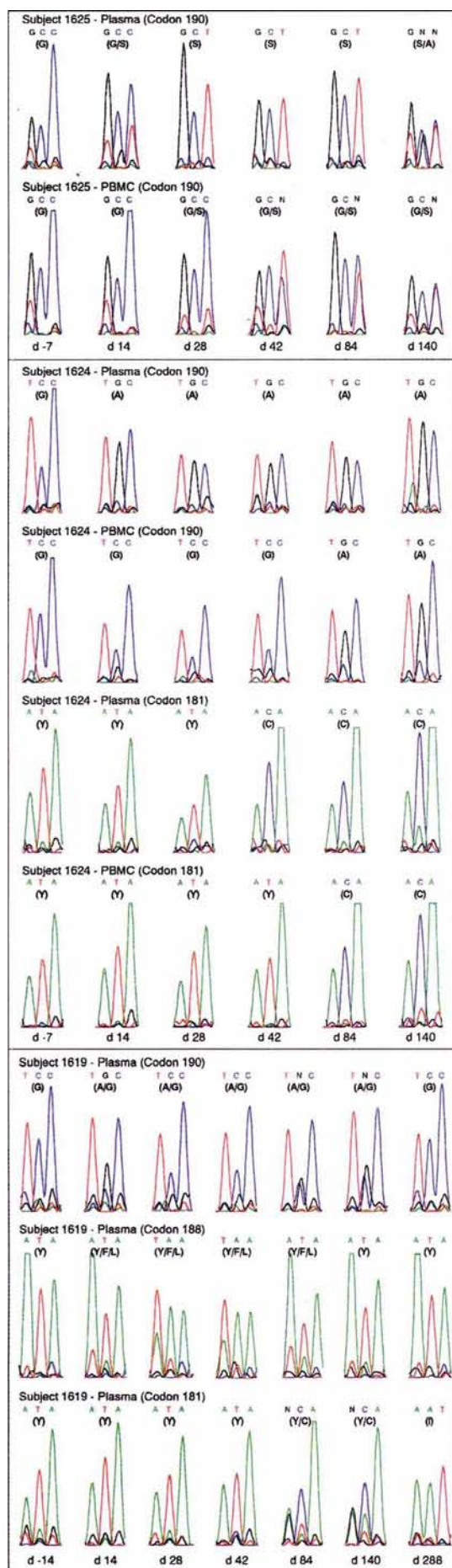
fluorescence units were as follows (predicted/observed): 0/<10%; 10/18%; 25/29%; 50/49%; 75/71% and 100/94%.

We next determined the ability of direct population sequencing to quantify wild-type and mutant viral RNA genomes in clinical specimens. Figure 3b shows the sequence chromatograms of RT codons 179–191 from virions pelleted directly from uncultured plasma specimens of subject 1625 before (day -7) and after (days +28 and +140) the initiation of NVP therapy. At day -7, all codons within the amino-terminal half of the RT gene (codons 1–250), including those shown, were wild-type at positions associated with NVP resistance^{31,32}. However, after only 28 days of NVP therapy, the wild-type plasma virus population was completely replaced by a NVP-resistant mutant population differing from the wild-type at codon 190 (glycine-to-serine substitution). After 140 days of drug therapy, this codon had evolved further such that the plasma virus population consisted of an equal mixture of two drug-resistant strains, one containing G190S and the other containing G190A. There were no other NVP-resistance-conferring mutations detectable within the viral RT gene.

FIG. 3 Quantitative detection of HIV-1 drug-resistance mutations by automated DNA sequencing. **a**, DNA sequence chromatograms of RT codon 190 from a defined mixture of wild-type (wt) and mutant (mut) HIV-1 cDNA clones differing only at the second base position of the codon. Sequences shown were obtained from, and therefore are presented as, the minus (non-coding) DNA strand. For example, the minus-strand TCC sequence shown corresponds to the plus-strand codon GGA (glycine, G). Similarly, the minus-strand TGC sequence corresponds to the plus-strand codon GCA (alanine, A). The single-letter amino-acid code corresponds to the plus-strand DNA sequence. Mixed bases approximating a 50/50 ratio are denoted as N. **b**, DNA sequence chromatograms of RT codons 179–191 (again displayed as the minus-strand sequence) derived from plasma-virion-associated RNA of subject 1625 before (day -7) and after (days +28 and +140) starting NVP therapy. Codon changes resulting in amino-acid substitutions at position 190 are indicated for the plus strand. For example, the GCC minus-strand sequence at position 190 (day -7) corresponds to GGC (glycine, G), and the GCT minus-strand sequence at position 190 (day +28) corresponds to AGC (serine, S) in the respective plus strands.

METHODS. Mixtures of wild-type and mutant cDNA clones (**a**) were prepared and diluted such that first-round PCR amplifications were done with 1,000 viral cDNA target molecules per reaction. HIV-1 RNA was isolated from virions pelleted from uncultured plasma specimens (**b**), as described¹⁸. cDNA was prepared using Moloney murine leukaemia virus reverse transcriptase (GIBCO BRL)⁶ and an oligonucleotide primer corresponding to nucleotides 4,283 to 4,302 of the HXB2 sequence⁴³. The full-length viral reverse transcriptase gene (1,680 bp) was amplified by means of a nested PCR using conditions and oligonucleotide primers (outer primers: nt 2,483–2,502 and 4,283–4,302; inner primers: nt 2,549–2,565 and 4,211–4,229), previously reported³⁰. Subgenomic fragments of the RT gene were also amplified using combinations of the following oligonucleotide primers: (5') 2,585–2,610; (5') 2,712–2,733; (3') 2,822–2,844; (3') 3,005–3,028; (3') 3,206–3,228; (3') 3,299–3,324; (3') 3,331–3,350; (3') 3,552–3,572; and (3') 3,904–3,921. All 3' primers incorporated the universal primer sequence for subsequent dye-primer sequence analysis. The HIV-1 copy number in every PCR reaction was determined (100–10,000 copies). A total of three to six separate PCR amplifications of primary patient material was done on each sample using different combinations of primers, and representative chromatograms are shown. Rarely, codon interpretation was ambiguous. In the day +140 plasma sample from subject 1625 (bottom of panel **b**), the complementary (plus) strand could read: AGC(serine), GCN(alanine), ACN(threonine), AGA/AGG(arginine) or GGN(glycine). In this case, we sequenced 7 full-length RT molecular clones and found that they encoded only serine or alanine. For sequencing, an automated ABI 373A sequencer and the Taq Dye Primer Cycle Sequencing Kit (ABI) were used. Sequences were analysed using Sequencher (Gene Codes Corp.) and Microgenie (Beckman) software packages, and base-pair mixtures were quantified by measuring relative peak-on-peak heights²⁸.





In all four subjects evaluated by direct viral population sequencing (Fig. 4), specific NVP-resistance-conferring mutations within the RT gene could be unambiguously identified and subsequently confirmed by molecular cloning, expression and drug susceptibility testing. In all cases, mutant virus increased rapidly in the plasma and virtually replaced wild-type virus after only 2–4 weeks of NVP therapy (Fig. 2c). By analysing the rate of accumulation of resistant mutants in the plasma population, we could obtain an independent estimate of the turnover rate of free virus. The rise of drug-resistant mutant virus is influenced substantially by the preceding increase in the CD4⁺ cell population (which provides additional resources for virus production³³) and therefore follows complex dynamics. However, we could obtain an estimate of these dynamics by making simplifying assumptions. We assume that wild-type virus declines exponentially with a decay rate α , and that the drug-resistant mutant increases exponentially with the rate β . Thus, the ratio of mutant to wild-type virus increases exponentially at the combined rate $\alpha + \beta$. Our genetic RNA (cDNA) data allow us to estimate this sum. Knowing α from our data on virus decline, we get $\beta \approx 0.27$, or a 32% daily virus production (average over 4 patients). Assuming that mutant virus rises exponentially, this corresponds to a doubling time of ~ 2 days, which is in excellent agreement with the measured elimination half-life of 2.0 ± 0.9 days for plasma virus (Figs 1 and 2a). Turnover of viral DNA from wild-type to drug-resistant mutant in PBMCs was delayed and less complete compared to plasma virus, reaching levels of only 50–80% of the total PBMC-associated viral DNA population by week 20 (Fig. 2c). Measurement of the time required for resistant virus to spread in the PBMC population allowed us also to estimate the half-life of infected PBMCs. After complete turnover of mutant virus in the plasma pool, we may assume that PBMCs infected with wild-type virus decline exponentially at a rate d , whereas cells infected by mutant virus are generated at a constant rate, but also decline exponentially at rate d . With these simplifying assumptions, the rate at which the frequency of resistant virus in the PBMC population increases provides an estimate for the parameter d and hence for the half-life of infected PBMCs. We obtained a half-life of ~ 50 –100 days. This means that the average half-life of infected PBMCs is very long and of the same order of magnitude as the half-life of uninfected PBMCs^{34,35}. Based on the long half-life of PBMCs, and the fact that these cells harbour predominantly wild-type virus at a time (days 14–28) when most virus in plasma is mutant, we conclude that most PBMCs contribute comparatively little to plasma virus load. Instead, other cell populations, most probably in the lymphoreticular system^{11,19,20}, must be the major source of virus production.

Direct sequence analysis of viral nucleic acid revealed not only rapid initial turnover in viral populations but also continuing viral evolution with respect to drug resistance mutations. In subject 1625 (Fig. 4, top panel), wild-type virus in plasma was completely replaced after 28 days of NVP therapy by mutant virus

FIG. 4 Quantitative detection of HIV-1 drug resistance mutations by automated DNA sequencing in plasma viral RNA (cDNA) and PBMC-associated viral DNA populations before and after the initiation of NVP on day 0. As in Fig. 3, minus-strand sequences are shown together with single-letter amino-acid codes of the corresponding plus-strand sequence. Mixed bases approximating a 50/50 ratio are denoted as N.

METHODS. HIV-1 cDNA was prepared from virions pelleted from uncultured plasma as described for Fig. 3. Viral DNA was isolated from uncultured PBMCs, as described⁴⁴. The full-length viral reverse transcriptase genes as well as subgenomic fragments were amplified and sequenced as described for Fig. 3. The HIV-1 copy number in every PCR reaction was determined (100–10,000 copies). Some sequences were determined from both coding and non-coding DNA strands to ensure the accuracy of quantitative measurements.

TABLE 1 *In situ* functional analysis of HIV-1 RT clones

Subject	Specimen		Functional clones	NVP-sensitive clones	NVP-resistant clones
1625	Plasma	day -7	80	80 (100%)	0 (0%)
		+14	72	27 (38%)	45 (62%)
		+28	57	0 (0%)	57 (100%)
		+84	67	0 (0%)	67 (100%)
		+140	86	0 (0%)	86 (100%)
1625	PBMC	-7	163	163 (100%)	0 (0%)
		+14	121	121 (100%)	0 (0%)
		+28	258	134 (52%)	124 (48%)
		+84	133	43 (32%)	90 (68%)
		+140	261	65 (25%)	196 (75%)
1624	Plasma	-7	19	19 (100%)	0 (0%)
		+14	34	4 (12%)	30 (88%)
		+28	79	6 (8%)	73 (92%)
		+140	27	0 (0%)	27 (100%)
1624	PBMC	-7	24	24 (100%)	0 (0%)
		+14	34	29 (85%)	5 (15%)
		+28	52	42 (81%)	10 (19%)
		+140	87	26 (30%)	61 (70%)
1605	PBMC	-7	31	31 (100%)	0 (0%)
		+140	31	11 (35%)	20 (65%)
1619	Plasma	-14	79	79 (100%)	0 (0%)
		+28	41	0 (0%)	41 (100%)
		+140	38	0 (0%)	38 (100%)

Full-length RT genes were amplified by PCR from uncultured plasma and uncultured PBMCs as described in Fig. 3 legend. DNA products were cloned into the *EcoRI* and *HindIII* sites of the bacterial expression plasmid pLG18-1 (refs 29, 30). The expression plasmids were screened for the presence of functional RT and tested *in situ* for susceptibility to NVP inhibition at 3,000 nM (~50–75 fold greater than the IC_{50})^{24,31,32}. To ensure accuracy in distinguishing RT genes encoding NVP-resistant versus sensitive enzymes, and to confirm the identification of specific NVP-resistance-conferring RT mutations obtained by direct sequencing (Figs 3 and 4), we determined the complete nucleotide sequences of 21 cloned RT genes which had been phenotyped in the *in situ* assay (V.A.J. and G.M.S., submitted). There was complete concordance between the phenotypes and genotypes of these 21 clones with respect to NVP-resistance-conferring mutations, as well as complete concordance between direct viral population sequences and clone-derived sequences at NVP-resistance-conferring codons.

(G190S), which in turn evolved by day 140 into a mixture of G190S and G190A. In subject 1624 (Fig. 4, middle panel), two codon changes conferring NVP resistance occurred. A G190A substitution appeared in plasma virus at day 14 and a Y181C appeared at day 42. Similarly, in subject 1605 (not shown), a Y181C mutation appeared in plasma at day 14 and a Y188L mutation at day 28. The sequential changes in plasma virus were mirrored by similar changes in PBMCs at later timepoints. In subject 1619, the pattern of resistance changes was even more complex (Fig. 4, bottom panel). By day 14, approximately 70% of plasma virus contained a G190A mutation. By day +28, this mutant population was largely replaced by virus containing a Y188F/L substitution. By day 84, still another major shift in the viral quasispecies occurred, this time resulting in a population of viruses containing mutations at both Y181C and G190A. Finally, by day 288 the viral population in plasma consisted exclusively of a mutant exhibiting a single tyrosine-to-isoleucine substitution at position 181 (Y181I); mutations at codons 188 and 190 were not present in this virus population. All of these amino-acid substitutions at RT codons 181, 188 and 190 were shown in our *in situ* expression studies and by others^{31,32,36} to confer high-level NVP resistance. The direct sequence analyses thus demonstrate that major changes in the HIV-1 quasispecies occur quickly and continuously in response to selection pressures and that these changes are reflected first and most prominently in the plasma virus compartment.

***In situ* RT gene expression and drug susceptibility testing.**

Because direct sequence analysis of viral mixtures provides only semiquantitative information and does not distinguish between viruses with functional rather than defective RT genes, we employed another method for quantifying virus turnover in uncultured plasma and PBMC compartments. Full-length RT genes were amplified by polymerase chain reaction (PCR),

cloned into pLG18-1, expressed in *Escherichia coli*, and tested individually for enzymatic function and NVP susceptibility by *in situ* assay^{29,30} (Table 1). For subject 1625 at day -7, 100% (80/80) of RT clones from plasma and 100% (163/163) of RT clones from PBMCs expressed enzyme that was sensitive to NVP inhibition. By day 14, however, 62% of plasma-derived clones expressed enzyme that was resistant to NVP, and by days 28, 84 and 140, 100% were resistant. Conversely, at day 14, 0% of PBMC-derived clones expressed NVP-resistant enzyme, and even after 28, 84 and 140 days, only 48–75% of clones were resistant. Similar results were obtained for the other study subjects (Table 1). Thus, the kinetics of virus population turnover determined by a quantitative RT *in situ* expression assay corresponded closely with those determined by direct population sequencing (Fig. 2c).

Infectious virus drug susceptibility testing. Plasma and PBMCs are known to harbour substantial proportions of defective or otherwise non-infectious virus^{6,37}. To determine whether the viral genomes represented in total viral nucleic acid (Fig. 4 and Table 1) corresponded to infectious virus with respect to NVP-resistance-conferring mutations, we co-cultivated PBMCs from three of the study subjects (1605, 1624, 1625) with normal donor lymphoblasts in order to establish primary virus isolates. The RT genes of these cultured viruses, obtained before and after therapy, were cloned (Fig. 3 and Table 1 legends) and sequenced in their entirety (V.A.J. and G.M.S., submitted). RT codons associated with NVP susceptibility were completely concordant in cultured and uncultured virus strains. Furthermore, the virus isolates exhibited NVP susceptibility profiles³⁸ consistent with their genotypes.

CD4⁺ lymphocyte dynamics

Changes in CD4⁺ lymphocyte counts during the first 28 days of therapy could be assessed in 17 of our patients (Fig. 2b and data not shown). CD4⁺ cell numbers increased in every patient by between 41 and 830 cells per mm³. For the entire group, the average increase was 186 ± 199 cells per mm³ (mean $\pm$ s.d.), or $268 \pm 319\%$ from baseline. As CD4⁺ lymphocytes increase in numbers because of (1) exponential proliferation of CD4⁺ cells in peripheral tissue compartments, and/or (2) constant (linear) production of CD4⁺ cells from a pool of precursors, we analysed our data based on each of these assumptions. The average percentage increase in cell number per day (assumption (1)) was $5.0 \pm 3.1\%$ (mean $\pm$ 1 s.d.). The average absolute increase in cell number per day (assumption (2)) was 8.0 ± 7.8 cells mm⁻³ d⁻¹. Given that peripheral blood contains only 2% of the total body lymphocytes³⁵ and that the average total blood volume is ~5 litres, an increase of 8 cells mm⁻³ d⁻¹ implies an overall steady-state CD4⁺ cell turnover rate (where increases equal losses) of $(50) \times (5 \times 10^6 \text{ mm}^3) \times (8 \text{ cells mm}^{-3} \text{ d}^{-1})$, or 2×10^9 CD4⁺ cells produced and destroyed each day.

Discussion

Previously, it was shown that lymphoreticular tissues serve as the primary reservoir and site of replication for HIV-1 (refs 11, 19, 20) and that virtually all HIV-1-infected individuals, regardless of clinical stage, exhibit persistent plasma viraemia in the range of 10^2 to 10^7 virions per ml⁶. However, the dynamic contributions of virus production and clearance, and of CD4⁺ cell infection and turnover, to the clinical 'steady-state' were obscure, although not unanticipated^{22,23,39}. We show by virus quantitation and mutation fixation rates that the composite lifespan of plasma virus and of virus-producing cells is remarkably short ($t_{1/2} = 2.0 \pm 0.9$ days). This holds true for patients with CD4⁺ lymphocyte counts as low as 18 cells per mm³ and as high as 355 cells per mm³ (Figs 1 and 2; G.M.S., unpublished). These findings were made in patients treated with three different antiretroviral agents having two entirely different mechanisms of action and using three different experimental approaches for assessing virus turnover. The viral kinetics thus cannot be

explained by a unique or unforeseen drug effect or a peculiarity of any particular virological assay method. Moreover, when new cycles of infection are interrupted by potent antiretroviral therapy, plasma virus levels fall abruptly by an average of 99%, and in some cases by as much as 99.99% (10,000-fold). This result indicates that the vast majority of circulating plasma virus derives from continuous rounds of *de novo* virus infection, replication and cell turnover, and not from cells that produce virus chronically or are latently infected and become activated. The identity and location of this actively replicating cell population is not known, but appears not to reside in the PBMC pool, consistent with prior reports^{11,19,20}. Nevertheless, PBMCs traffic through secondary lymphoid organs and to some extent are in equilibrium with these cells³⁵. It is thus possible that a small fraction of PBMCs^{8,9,14-17}, like a small fraction of activated lymphoreticular cells²⁰, could make an important contribution to viraemia.

The magnitude of ongoing virus infection and production required to sustain steady-state levels of viraemia is extraordinary: based on a virus $t_{1/2}$ of 2.0 days and first-order clearance kinetics ($v(t) = v(0)e^{-\alpha t}$, where $\alpha = 0.693/t_{1/2}$), 30% or more of the total virus population in plasma must be replenished daily. For a typical HIV-1-infected individual with a plasma virus titre equalling the pretreatment geometric mean in this study ($10^{5.5}$ RNA molecules per ml/2 RNA molecules per virion = $10^{5.2}$ virions per ml) and a plasma volume of 3 litres, this amounts to $(0.30) \times (10^{5.2}) \times (3 \times 10^3) = 1.1 \times 10^8$ virions per day (range for all 22 subjects, 2×10^7 to 7×10^9). Even this may be a substantial underestimate of virus expression because virions may be inefficiently transported from the interstitial extravascular spaces into the plasma compartment and viral protein expression alone (short of mature particle formation) may result in cytopathy or immune-mediated destruction. Because the half-life of cells producing the majority of plasma virus cannot exceed 2.0 days, at least 30% of these cells must also be replaced daily. In our patients, we estimated the rate of CD4⁺ lymphocyte turnover to be, on average, 2×10^9 cells per day, or about 5% of the total CD4⁺ lymphocyte population, depending on clinical stage. This rapid and ongoing recruitment of CD4⁺ cells into a short-lived virus-expressing pool probably explains the abrupt increase in CD4⁺ lymphocyte numbers that is observed immediately following the initiation of potent antiretroviral therapy, and suggests the possibility of successful immunological reconstitution even

in late-stage disease if effective control of viral replication can be sustained.

The kinetics of virus and CD4⁺ lymphocyte production and clearance reported here have a number of biological and clinical implications. First, they are indicative of a dynamic process involving continuous rounds of *de novo* virus infection, replication and rapid cell turnover that probably represents a primary driving force underlying HIV-1 pathogenesis. Second, the demonstration of rapid and virtually complete replacement of wild-type virus by drug-resistant virus in plasma after only 14–28 days of drug therapy is a striking example of the capacity of the virus for biologically relevant change. In particular, this implies that HIV-1 must have enormous potential to evolve in response to selection pressures as exerted by the immune system³⁹. Although other studies^{40–42} have provided some evidence that virus turnover occurs sooner in plasma than in PBMCs, our data show this phenomenon most clearly. A similar experimental approach involving the genotypic and phenotypic analysis of plasma virus could be helpful in identifying viral mutations and selection pressures involved in resistance to other drugs, immune surveillance and viral pathogenicity. Third, the difference in lifespan between virus-producing cells and latently infected cells (PBMCs) suggests that virus expression *per se* is directly involved in CD4⁺ cell destruction. The data do not suggest an ‘innocent bystander’ mechanism of cell killing whereby uninfected or latently infected cells are indirectly targeted for destruction by adsorption of viral proteins or by autoimmune reactivities.

Although we have emphasized that most virus in plasma derives from an actively replicating short-lived population of cells, latently infected cells that become activated or chronically producing cells that generate proportionately less virus (and thus do not contribute substantially to the plasma virus pool) may nonetheless be important in HIV-1 pathogenesis. Based on *in situ* analysis²⁰, these cells far outnumber the actively replicating pool and the diversity of their constituent viral genomes represents a potentially important source of clinically relevant variants, including those conferring drug resistance. In future studies, it will be important not only to discern the specific elimination rates of free virus and of the most actively producing cells, but also the dynamics of virus replication and cell turnover in other cell populations and in patients at earlier stages of infection. Such information will be essential to developing a better understanding of HIV-1 pathogenesis and a more rational approach to therapeutic intervention. □

Received 22 November; accepted 16 December 1994.

- Ho, D. D., Moudgil, T. & Alam, M. *New Engl. J. Med.* **321**, 1621–1625 (1989).
- Coombs, R. W. et al. *New Engl. J. Med.* **321**, 1626–1631 (1989).
- Saag, M. S. et al. *J. Infect. Dis.* **164**, 72–80 (1991).
- Clark, S. J. et al. *New Engl. J. Med.* **324**, 954–960 (1991).
- Daar, E. S., Moudgil, T., Meyer, R. D. & Ho, D. D. *New Engl. J. Med.* **324**, 961–964 (1991).
- Piatak, M. Jr et al. *Science* **259**, 1749–1754 (1993).
- Piatak, M. et al. *Lancet* **341**, 1099 (1993).
- Schnittman, S. M., Greenhouse, J. J., Lane, H. C., Pierce, P. F. & Fauci, A. S. *AIDS Res. Hum. Retrovir.* **7**, 361–367 (1991).
- Michael, N. L., Vahey, M., Burke, D. S. & Redfield, R. R. *J. Virol.* **66**, 310–316 (1992).
- Winters, M. A., Tan, L. B., Katzenstein, D. A. & Merigan, T. C. *J. clin. Microbiol.* **31**, 2960–2966 (1993).
- Pantaleo, G. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9838–9842 (1991).
- Connor, R. I., Mohri, H., Cao, Y. & Ho, D. D. *J. Virol.* **67**, 1772–1777 (1993).
- Bagnarelli, P. et al. *J. Virol.* **66**, 7328–7335 (1992).
- Bagnarelli, P. et al. *J. Virol.* **68**, 2495–2502 (1994).
- Graziosi, C. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6405–6409 (1993).
- Patterson, B. K. et al. *Science* **260**, 976–979 (1993).
- Saksela, K., Stevens, C., Rubinstein, P. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **91**, 1104–1108 (1994).
- Cao, Y. et al. *AIDS Res. Hum. Retrovir.* (in the press).
- Pantaleo, G. et al. *Nature* **362**, 355–358 (1993).
- Embertson, J. et al. *Nature* **362**, 359–362 (1993).
- Aoki-Sei, S. et al. *AIDS Res. Hum. Retrovir.* **8**, 1263–1270 (1992).
- Coffin, J. M. *Curr. Top. Microbiol. Immun.* **176**, 143–164 (1992).
- Wain-Hobson, S. *Curr. Opin. Genet. Dev.* **3**, 878–883 (1993).
- Merluzzi, V. J. et al. *Science* **250**, 1411–1413 (1990).
- Kempf, D. et al. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

- Vacca, J. P. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4096–4100 (1994).
- Mulder, J. et al. *J. clin. Microbiol.* **32**, 292–300 (1994).
- Larder, B. A. et al. *Nature* **365**, 671–675 (1993).
- Prasad, V. R. & Goff, S. P. *J. biol. Chem.* **264**, 16689–16693 (1989).
- Saag, M. S. et al. *New Engl. J. Med.* **329**, 1065–1072 (1993).
- Richman, D. D. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 11241–11245 (1991).
- Richman, D. D. et al. *J. Virol.* **68**, 1660–1666 (1994).
- McLean, A. R. & Nowak, M. A. *AIDS* **6**, 71–79 (1992).
- Michie, C., McLean, A., Alcock, C. & Beverley, P. C. L. *Nature* **360**, 264–265 (1992).
- Sprent, J. & Tough, D. F. *Science* **265**, 1395–1400 (1994).
- Balzarini, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 6599–6603 (1994).
- Meyerhans, A. et al. *Cell* **58**, 901–910 (1989).
- Japour, A. J. et al. *Antimicrob. Agents Chemother.* **37**, 1095–1101 (1993).
- Nowak, M. A. et al. *Science* **254**, 963–969 (1991).
- Simmonds, P. et al. *J. Virol.* **65**, 6266–6276 (1991).
- Smith, M. S., Koerber, K. L. & Pagano, J. S. *J. infect. Dis.* **167**, 445–448 (1993).
- Zhang, Y.-M., Dawson, S. C., Landsman, D., Lane, H. C. & Salzman, N. P. *J. Virol.* **68**, 425–432 (1994).
- Myers, G. K., Korber, B., Berzofsky, J. A. & Smith, R. F. *Human Retroviruses and AIDS 1993* (Los Alamos National Laboratory, New Mexico, 1993).
- Shaw, G. M. et al. *Science* **226**, 1165–1171 (1984).

ACKNOWLEDGEMENTS. We thank the study participants; K. Squires, J. M. Kilby, M. Trechsel, L. DeLoach and the UAB 1917 Clinic staff; Abbott Laboratories, Merck & Co. and Boehringer Ingelheim Pharmaceuticals Inc. (BIPI); J. Coffin, R. May and F. Gao for discussion; J. Decker, S. Campbell-Hill, Y. Niu and S. Yin Jiang for technical assistance; and J. Wilson for artwork. This study was supported by the NIH, the US Army Medical Research Acquisition Activity, BIPI, the Wellcome Trust, Keble College and Boehringer Ingelheim Stiftung. Core research facilities were provided by the UAB Center for AIDS Research, the UAB AIDS Clinical Trials Unit and the Birmingham Veterans Administration Medical Center.

Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection

David D. Ho, Avidan U. Neumann^{*†}, Alan S. Perelson[†], Wen Chen, John M. Leonard[‡] & Martin Markowitz

Aaron Diamond AIDS Research Center, NYU School of Medicine, 455 First Avenue, New York, New York 10016, USA

^{*} Santa Fe Institute, Santa Fe, New Mexico 87501, USA

[†] Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

[‡] Pharmaceutical Products Division, Abbott Laboratories, Abbott Park, Illinois 60064, USA

Treatment of infected patients with ABT-538, an inhibitor of the protease of human immunodeficiency virus type 1 (HIV-1), causes plasma HIV-1 levels to decrease exponentially (mean half-life, 2.1 ± 0.4 days) and CD4 lymphocyte counts to rise substantially. Minimum estimates of HIV-1 production and clearance and of CD4 lymphocyte turnover indicate that replication of HIV-1 *in vivo* is continuous and highly productive, driving the rapid turnover of CD4 lymphocytes.

In HIV-1 pathogenesis, an increased viral load correlates with CD4 lymphocyte depletion and disease progression^{1,9}, but relatively little information is available on the kinetics of virus and CD4 lymphocyte turnover *in vivo*. Here we administer an inhibitor of HIV-1 protease, ABT-538 (refs 10, 11), to twenty infected patients in order to perturb the balance between virus production and clearance. From serial measurements of the subsequent changes in plasma viraemia and CD4 lymphocyte counts, we have been able to infer kinetic information about the pretreatment steady state.

ABT-538 has potent antiviral activity *in vitro* and favourable pharmacokinetic and safety profiles *in vivo*¹⁰. It was administered orally (600–1,200 mg per day) on day 1 and daily thereafter to twenty HIV-1-infected patients, whose pretreatment CD4 lymphocyte counts and plasma viral levels ranged from 36 to 490 per mm³ and from 15×10^3 to 554×10^3 virions per ml, respectively (Table 1). Post-treatment CD4 lymphocyte counts were monitored sequentially, as were copy numbers of particle-associated HIV-1 RNA in plasma, using an ultrasensitive assay (Fig. 1 legend) based on a modification of the branched DNA signal-amplification technique^{12,13}. The trial design and clinical findings of this study will be reported elsewhere (M.M. *et al.*, manuscript in preparation).

Kinetics of HIV-1 turnover

Following ABT-538 treatment, every patient had a rapid and dramatic decline in plasma viraemia over the first two weeks. As shown using three examples in Fig. 1a, the initial decline in plasma viraemia was always exponential, demonstrated by a straight-line fit to the data on a log plot. The slope of this line, as defined by linear regression, permitted the half-life ($t_{1/2}$) of viral decay in plasma to be determined (Fig. 1 legend): for example, patient 409 was found to have a viral decay slope of -0.47 per day, yielding a $t_{1/2}$ of 1.5 days (Fig. 1a). Hence the rate and extent of decay of plasma viraemia was determined for each patient. As summarized in Fig. 1b, in every case there was a rapid decline, the magnitude of which ranged from 11- to 275-fold, with a mean of 66-fold (equivalent to 98.5% inhibition). The residual viraemia may be attributable to inadequate drug concentration in certain tissues, drug resistance, persistence of a small long-lived virus-producing cell population (such as macrophages), and gradual activation of a latently infected pool of cells. As summarized in Table 1, the viral decay slopes varied from -0.21 to -0.54 per day, with a mean of -0.34 ± 0.06 per day; correspondingly, $t_{1/2}$ varied from 1.3 to 3.3 days, with a mean of 2.1 ± 0.4 days. The latter value indicates that, on aver-

age, half of the plasma virions turn over every two days, showing that HIV-1 replication *in vivo* must be highly productive.

The exponential decline in plasma viraemia following ABT-538 treatment reflects both the clearance of free virions and the loss of HIV-1-producing cells as the drug substantially blocks new rounds of infection. But although drug inhibition is probably incomplete and virus-producing cells are not lost immediately, a minimum value for viral clearance can still be determined (Fig. 1 legend) by multiplying the absolute value of the viral decay slope by the initial viral load. Assuming that ABT-538 administration does not affect viral clearance, this estimate is also valid before treatment. As the viral load varies little during the pretreatment phase (Fig. 1a, and data not shown), we assume there exists a steady state and hence the calculated clearance rate is equal to the minimum virion production rate before drug therapy. Factoring in the patient's estimated plasma and extracellular fluid volumes based on body weight, we determined the minimum daily production and clearance rate of HIV-1 particles for each case (Table 1). These values ranged from 0.05 to 2.07×10^9 virions per day with a mean of $0.68 \pm 0.13 \times 10^9$ virions per day. Although these viral turnover rates are already high, true values may be up to a few-fold higher, depending on the $t_{1/2}$ of virus-producing lymphocytes. The precise kinetics of this additional parameter remains undefined. However, the mean $t_{1/2}$ of virus-producing cells is probably less, or in any case cannot be much larger, than the mean $t_{1/2}$ of 2.1 days observed for plasma virion elimination, demonstrating that turnover of actively infected cells is both rapid and continuous. It could also be inferred from our data that nearly all (98.5%) of the plasma virus must come from recently infected cells.

Examination of Fig. 1b shows that the viral decay slopes (clearance rate constants) are independent of the initial viral loads. The slopes do not correlate with the initial CD4 lymphocyte counts (Fig. 2a), another indicator of the disease status of patients. Therefore these observations strongly suggest that the viral clearance rate constant is not dependent on the stage of HIV-1 infection. Instead, they indicate that viral load is largely a function of viral production, because clearance rate constants vary by about 2.5-fold whereas the initial loads vary by almost 40-fold (Table 1).

Kinetics of CD4 lymphocyte turnover

After ABT-538 treatment, CD4 lymphocyte counts rose in each of the 18 patients that could be evaluated. As shown in three examples in Fig. 3, some increases were dramatic (patient 409, for example) whereas others (such as patient 303) were modest. Based on the available data, it was not possible to determine

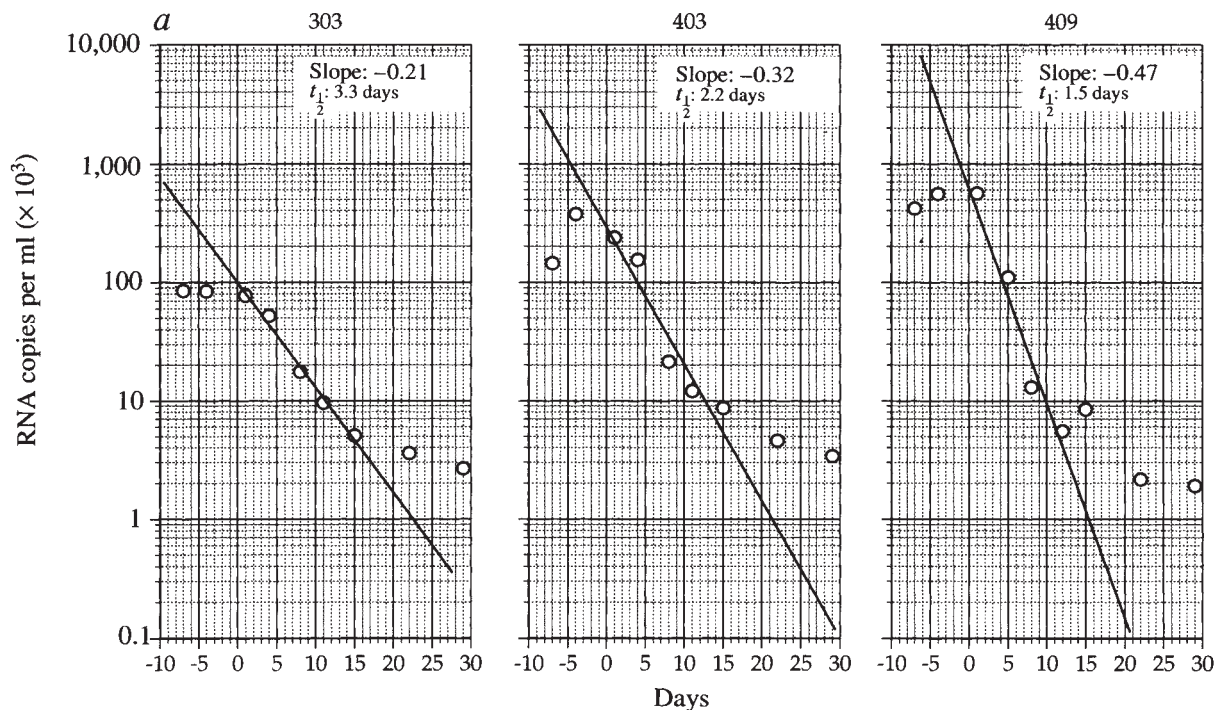
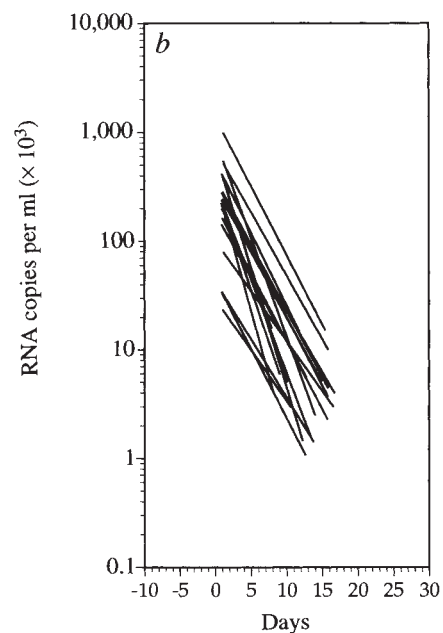


FIG. 1 *a*, Plasma viral load before and after ABT-538 treatment was begun on day 1 for three representative cases. Plasma samples were tested with the branched DNA signal-amplification assay as previously described^{12,13}. Those samples with RNA levels below the detection sensitivity of 10,000 copies per ml were then tested using a modified assay differing from the original in two ways: hybridization of the bDNA amplification system is mediated by binding to overhangs on contiguous target probes; and the enzymatic amplification system has been enhanced by modification of wash buffers and the solution in which the alkaline phosphatase probe is diluted. The results of these changes are a diminution of background signals, an enhancement of alkaline phosphatase activity, and thus a greater detection sensitivity (500 copies per ml). Linear regression was used to obtain the best-fitting straight line for 3–5 data points between day 1 and the inflection point before the plateau of the new steady-state level. The slope, S , of each line represents the rate of exponential decrease; that is, the straight-line fit indicates that the viral load decreases according to $V(t) = V(0) \exp(-St)$. Given the exponential decay, $V(t_{1/2}) = V(0)/2 = V(0) \exp(-St_{1/2})$, and hence the viral half life, $t_{1/2} = \ln(2)/S$. Before drug administration, the change in viral load with time can be expressed by the differential equation, $dV/dt = P - cV$, where P is the viral production rate, c is the viral clearance rate constant, and V is the number of plasma virions. During the pretreatment steady state, $dV/dt = 0$, and hence $P = cV$. We have also tested more intricate models, in which the viral decay is governed by two or three exponential rates, namely the viral clearance rate, the decay rate of virus-producing cells, and the decay rate of latently infected cells. But there are insufficient data at this time to estimate the multiple parameters separately. Nonetheless, using the model in ref. 14, we find that if the death rate of latently infected cells is very small compared with the other two rates, then viral decay follows a simple exponential decline, with $S = c$, because the slow activation of a large number of latently infected cells offsets the loss of actively infected cells. Irrespective of the model, on a log plot, $S = -d(\ln V)/dt = c - P/V$. If drug inhibition is complete and virus-

with confidence whether the rise was strictly exponential (Fig. 3, top) or linear (Fig. 3, bottom). An exponential increase would be consistent with proliferation of CD4 lymphocytes in the periphery, particularly in secondary lymphoid organs, whereas a linear increase would indicate cellular production from a precursor source such as the thymus¹⁴. Given that the thymus involutes with age and becomes further depleted with HIV-1 infection¹⁵, it is more likely that the rise in CD4 lymphocytes is largely due to proliferation. Nevertheless, as both components may contribute, we analysed the observed CD4 lymphocyte data by modelling both exponential and linear increases.



producing cells are rapidly lost (so $P = 0$), then $S = c$. If viral production continues, S is still $\leq c$, so the slope is a minimum estimate of the viral clearance efficiency. *b*, Decline of plasma viral load after ABT-538 treatment in all 20 patients. The slope for each case was obtained as already discussed, and the length of each line was determined by the initial viral load and the new steady-state level.

The slope of the line depicting the rise in CD4 lymphocyte counts on a log plot was determined for each case (Fig. 3, top). Individual slopes varied considerably, ranging from 0.004 to 0.088 per day, with a mean of 0.047 per day (Table 1), corresponding to a mean doubling time of ~ 15 days (Fig. 3 legend). On average, the entire population of peripheral CD4 lymphocytes was turning over every 15 days in our patients during the pretreatment steady state when CD4 lymphocyte production and destruction were balanced. Moreover, the slopes were inversely correlated with baseline CD4 lymphocyte counts (Fig. 2*b*) in that patients with lower initial CD4 cell counts had more pro-

TABLE 1 Summary data of HIV-1 and CD4 lymphocyte turnover during the pretreatment steady state

Baseline values			Kinetics of HIV-1 turnover			Kinetics of CD4 lymphocyte turnover*		
Patient	CD4 cell count (mm ⁻³)	Plasma viraemia (virions per ml × 10 ³)†	Slope	t _{1/2} (days)	Minimum production and clearance‡ (virions per day × 10 ⁹)	Minimum production and destruction		
						Slope	Blood§ (cells per day × 10 ⁶)	Total (cells per day × 10 ⁹)
301	76	193	-0.30	2.3	0.56	0.070 (6.9)	21.7 (28.1)	1.1 (1.4)
302	209	80	-0.27	2.6	0.26	0.004 (0.5)	4.3 (2.7)	0.2 (0.1)
303	293	41	-0.21	3.3	0.11	0.005 (1.4)	9.9 (9.5)	0.5 (0.5)
304	174	121	-0.28	2.5	0.54	0.019 (1.9)	22.2 (13.0)	1.1 (0.6)
305	269	88	-0.33	2.1	0.50	0.055 (21.5)	108.0 (157.0)	5.4 (7.8)
306	312	175	-0.52	1.3	1.27	0.058 (25.7)	105.0 (150.0)	5.3 (7.5)
308	386	185	-0.46	1.5	1.48	0.020 (9.1)	55.9 (65.8)	2.8 (3.3)
309	49	554	-0.29	2.4	1.85	0.088 (11.8)	20.7 (56.6)	1.0 (2.8)
310	357	15	-0.26	2.7	0.05	0.038 (15.6)	71.0 (81.9)	3.6 (4.1)
311	107	130	-0.29	2.4	0.51	0.064 (11.0)	38.9 (62.8)	2.0 (3.1)
312	59	70	-0.30	2.3	0.30	0.048 (4.5)	17.0 (26.9)	0.8 (1.4)
313	47	100	-0.54	1.3	0.88	0.077 (5.9)	24.7 (40.5)	1.2 (2.0)
401	228	101	-0.40	1.7	0.47	NA	NA	NA
402	169	55	-0.28	2.5	0.21	0.014 (3.1)	13.4 (17.4)	0.7 (0.9)
403	120	126	-0.32	2.2	0.74	0.015 (2.4)	13.8 (18.7)	0.7 (0.9)
404	46	244	-0.27	2.6	1.06	0.080 (8.5)	24.6 (57.5)	1.2 (2.9)
406	490	18	-0.31	2.2	0.08	NA	NA	NA
408	36	23	-0.25	2.8	0.08	0.059 (3.4)	12.5 (19.7)	0.6 (1.0)
409	67	256	-0.47	1.5	2.07	0.073 (15.9)	35.3 (115.0)	1.8 (5.7)
410	103	99	-0.36	1.9	0.53	0.051 (5.6)	32.4 (34.5)	1.6 (1.7)
			-0.21 to					
Range	36-490	15-554	-0.54	1.3-3.3	0.05-2.07	0.004-0.088 (0.5-25.7)	4.3-108.0 (2.7-157.0)	0.2-5.4 (0.1-7.8)
Mean	180 ± 46	134 ± 40	-0.34 ± 0.06	2.1 ± 0.4	0.68 ± 0.13	0.047 (8.6)	35.1 (53.2)	1.8 (2.6)

* The results for the kinetics of CD4 lymphocyte turnover generated by an exponential growth model are shown without parentheses; results generated by a linear production model are shown in parentheses.

† Each virion contains two RNA copies.

‡ Calculated using plasma and extracellular fluid volumes estimated from body weights, and assuming that plasma and extracellular fluid compartments are in equilibrium.

§ Calculated using blood volumes estimated from body weights.

|| Calculated on the assumption that the lymphocyte pool in blood represents 2% of the total population¹⁶. NA, not analysed owing to large fluctuations in CD4 cell counts.

minent rises. This demonstrates convincingly that the CD4 lymphocyte depletion seen in AIDS is primarily a consequence of the destruction of these cells induced by HIV-1 not a lack of their production.

As ABT-538 treatment reduces virus-mediated destruction of CD4 lymphocytes, the observed increase in CD4 cells provides a minimum estimate (Fig. 3 legend) of the pretreatment CD4 lymphocyte production rate, which in turn equals the destruction rate during the steady state. Minimum production (destruction)

rates were calculated for each case by multiplying the slope, the initial CD4 cell count, and the estimated blood volume. The minimum numbers of CD4 cells in blood produced or destroyed each day ranged from 4.3×10^6 to 108×10^6 , with a mean of 35.1×10^6 (Table 1). Given that the blood lymphocyte pool is about 2% of the total population¹⁶, the overall CD4 lymphocyte turnover in our patients was calculated to vary from 0.2×10^9 to 5.4×10^9 cells per day, with a mean of 1.8×10^9 cells per day.

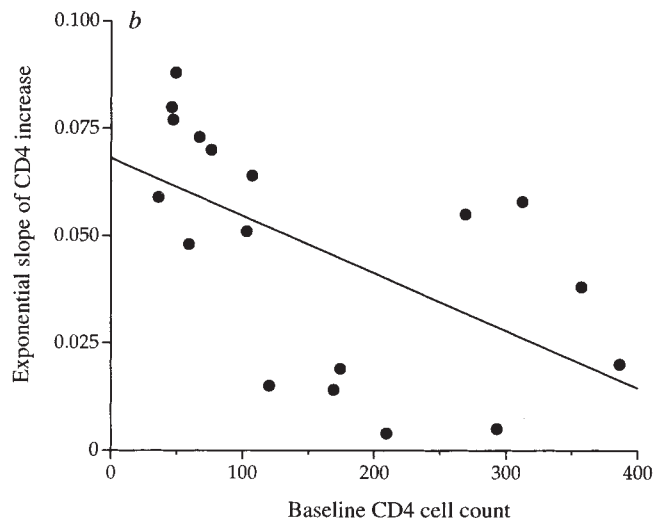
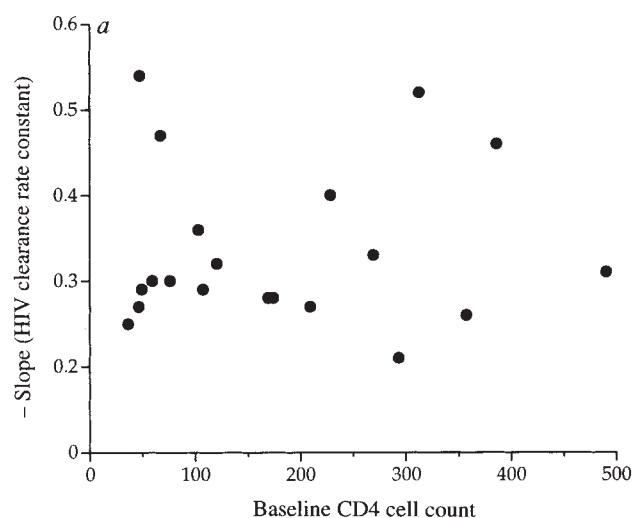
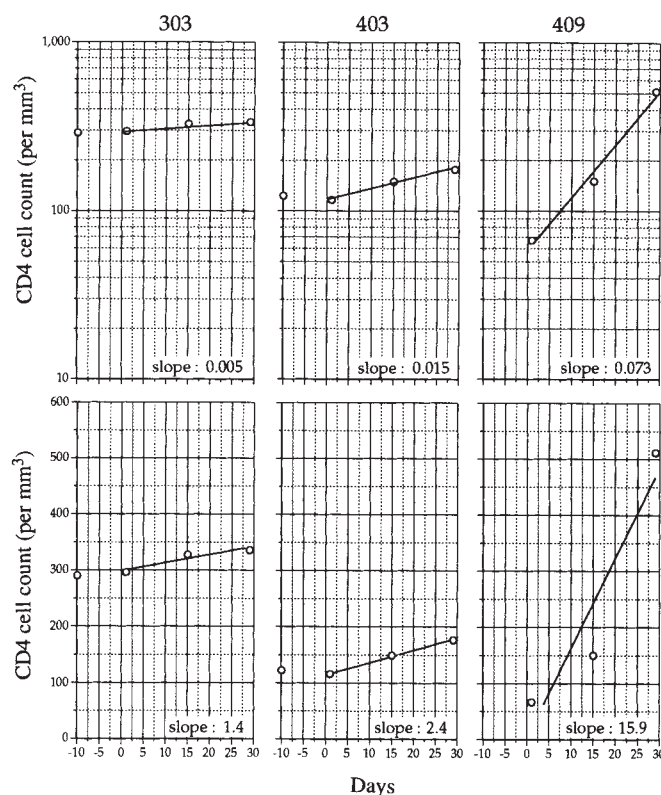


FIG. 2 a, Lack of correlation between viral decay slopes and disease status as indicated by baseline CD4 cell counts. Correlation coefficient = 0.05 (P value > 0.1). b, Inverse correlation between the exponential CD4 increase slopes and baseline CD4 cell counts. Correlation coefficient = -0.57 (P value < 0.01). Such an inverse correlation

would be expected if T-cell proliferation were governed by a density-dependent growth function (logistic, for example), in which the growth rate decreases with increasing population level, if T cells were produced from precursors at a constant rate or from a combination of these two effects.

FIG. 3 Increase in CD4 cell counts after ABT-538 treatment plotted on a logarithmic (top) or linear (bottom) scale. Each slope was obtained from the best-fit line derived from linear regression on 2–4 data points. In the model for exponential increase, the doubling time was determined by dividing $\ln(2)$ by the slope. From the slope, we obtained minimum estimates of the CD4 lymphocyte production rate. The change in CD4 cell number over time can be described by the equation, $dT/dt = P - \mu T$, where T is the cell count, P is the cell production rate, and μ is the cell decay rate. The slope, S , on a log plot is thus $= d(\ln T)/dt = P/T - \mu$. Hence, $S \times T$ must be less than P , showing that our estimates indeed represent minimum CD4 lymphocyte production rates. Using a similar argument, slopes derived from a model of linear increases are also minimum estimates of CD4 lymphocyte production.



The increase in CD4 lymphocyte counts following ABT-538 administration was also modelled linearly (Fig. 3, bottom). The slope of the line depicting the increase for each case was determined, and the values varied from 0.5 to 25.7 cells per mm^3 per day, with a mean of 8.6 cells per mm^3 per day (Table 1). Using the same argument as for the exponential case, minimum estimates of total CD4 production (or destruction) rates at baseline were determined to vary from 0.1×10^9 to 7.8×10^9 cells per day, with a mean of 2.6×10^9 cells per day.

Although our two sets of CD4 lymphocyte analyses do not yield identical numerical results, they are in close agreement and emphasize the same qualitative points about HIV-1 pathogenesis. The number of CD4 lymphocyte destroyed and replenished each day is of the order of 10^9 , which is strikingly close to estimates of the total number of HIV-1 RNA-expressing lymphocytes in the body determined using *in situ* polymerase chain reaction and hybridization methods^{8,17}. In addition, CD4 replenishment appears to be highly stressed in many patients in that the faster production rates are ~25–78-fold higher than the slowest rate (Table 1), which is presumably still higher than the as-yet-undefined normal CD4 turnover rate. The precise mechanisms of CD4 lymphocyte repopulation, however, will have to be addressed in the future by studies on phenotypic markers and functional status of the regenerating cells. Nonetheless, the rapid CD4 lymphocyte turnover has several implications. First, the apoptosis commonly observed in the setting of HIV-1 infection¹⁸ may simply be an expected consequence of an active lymphocyte

regenerative process. Second, the CD4 lymphocyte depletion seen in advanced HIV-1 infection may be likened to a sink containing a low water level, with the tap and drain both equally wide open. As the regenerative capacity of the immune system is not infinite, it is not difficult to see why the sink eventually empties. It is also evident from this analogy that our primary strategy to reverse the immunodeficiency ought to be to target virally mediated destruction (plug the drain) rather than to emphasize lymphocyte reconstitution (put in a second tap).

Discussion

We believe our new kinetic data have important implications for HIV-1 therapy and pathogenesis. It is self evident that, with rapid turnover of HIV-1, generation of viral diversity and the attendant increased opportunities for viral escape from therapeutic agents are unavoidable sequelae^{19,20}. Treatment strategies, if they are to have a dramatic clinical impact, must therefore be initiated as early in the infection course as possible, perhaps even during seroconversion. The rapid turnover of HIV-1 in plasma also suggests that current protocols for monitoring the acute antiviral activity of novel compounds must be modified to focus on the first few days following drug initiation. Our interventional approach to AIDS pathogenesis has shown that HIV-1 production and clearance are delicately balanced but highly dynamic processes. Taken together, our findings strongly support the view that AIDS is primarily a consequence of continuous, high-level replication of HIV-1, leading to virus- and immune-mediated killing of CD4 lymphocytes. □

Received 16 November; accepted 15 December 1994.

- Weiss, R. A. *Science* **260**, 1273–1279 (1993).
- Ho, D. D., Moudgil, T. & Alam, M. *New Engl. J. Med.* **321**, 1621–1625 (1989).
- Simmonds, P. et al. *J. Virol.* **64**, 864–872 (1990).
- Connor, R. I., Mohri, H., Cao, Y. & Ho, D. D. *J. Virol.* **67**, 1772–1777 (1993).
- Patterson, B. K. et al. *Science* **260**, 976–979 (1993).
- Platak, M. et al. *Science* **259**, 1749–1754 (1993).
- Pantaleo, G. et al. *Nature* **362**, 355–359 (1993).
- Embertson, J. et al. *Nature* **362**, 359–362 (1993).
- Wain-Hobson, S. *Nature* **366**, 22 (1993).
- Kempf, D. et al. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Markowitz, M. et al. *J. Virol.* (in the press).
- Pachl, C. et al. *J. AIDS* (in the press).
- Cao, Y. et al. *AIDS Res. Human Retroviruses* (in the press).

- Perelson, A. S., Kirschner, D. E. & De Boer, R. J. *Math. Biosci.* **114**, 81–125 (1993).
- Grody, W. W., Fligiel, S. & Naeim, F. *Am. J. clin. Path.* **84**, 85–95 (1985).
- Paul, W. E. *Fundamental Immunology* (Raven, New York, 1985).
- Haase, A. T. *Ann. N. Y. Acad. Sci.* **724**, 75–86 (1994).
- Ameisen, J. C. & Capron, A. *Immun. Today* **12**, 102–105 (1991).
- Wain-Hobson, S. *Curr. Opin. Genet. Dev.* **3**, 878–883 (1993).
- Coffin, J. M. *Curr. Top. Microbiol. Immun.* **176**, 143–164 (1992).

ACKNOWLEDGEMENTS. We thank the patients for their participation, Y. Cao and J. Wilbur for assistance with branched DNA assays, A. Hsu and J. Valdes for input on trial design, and J. Moore and R. Koup for helpful discussions. This work was supported by grants from Abbott Laboratories, the NIH and NYU Center for AIDS Research, the Joseph P. and Jeanne M. Sullivan Foundation, Los Alamos National Laboratory LRDR Program and The Aaron Diamond Foundation.

Evidence for a black hole from high rotation velocities in a sub-parsec region of NGC4258

Makoto Miyoshi*, James Moran†, James Herrnstein†, Lincoln Greenhill†, Naomasa Nakai‡, Phillip Diamond§ & Makoto Inoue‡

* Mizusawa Astrogeodynamics Observatory, National Astronomical Observatory, 2-12 Hoshigaoka, Mizusawa Iwate 023, Japan

† Harvard-Smithsonian Center for Astrophysics, Mail Stop 42, 60 Garden Street, Cambridge, Massachusetts 02138, USA

‡ Nobeyama Radio Observatory, National Astronomical Observatory, Minamisaku, Minamimaki, Nagano 384-13, Japan

§ National Radio Astronomy Observatory, PO Box 0, Socorro, New Mexico 87801, USA

MANY galaxies are thought to contain massive black holes—exceeding ten million solar masses—at their centres^{1,2}, but firm observational evidence has proved to be surprisingly elusive. The best evidence comes from observing gas or stars rotating rapidly within a small region around a central body. If the observed velocities are due solely to the gravitational force of the central body—as in the Solar System—then the mass of the central body can be readily calculated. Here we present observations of rotating gas near the centre of the galaxy NGC4258 (M106), which indicate the presence of a mass of 3.6×10^7 solar masses in a region less than 0.13 pc in radius. The volume-averaged mass density in this region exceeds by a factor of at least 40 that for any other black-hole candidate observed previously. These observations provide compelling evidence that a massive black hole exists at the centre of NGC4258.

We observed line emission from water masers in a region surrounding the centre of NGC4258. (Masers are the microwave equivalent of lasers.) These emission lines are very useful for studies of this kind, because they are both strong and intrinsically narrow. Powerful water maser emission has been previously detected at the centre of this galaxy by Claussen *et al.*³, who discussed the possible existence of a circumnuclear disk⁴.

Nakai *et al.*⁵ found extremely high-velocity features symmetrically offset from the systemic velocity by $\sim 1,000 \text{ km s}^{-1}$ in radial velocity. Previous very-long-baseline interferometry (VLBI) observations⁶ and time series analysis of the spectra^{7,8} suggest that the masers originate in a rotating disk surrounding a massive object (with mass M of 1.5×10^7 solar masses ($M_\odot$)) at the centre of the galaxy⁹.

We observed NGC4258 for 14 h on 26 April 1994 using the Very Long Baseline Array (VLBA) of the National Radio Astronomy Observatory augmented by the Very Large Array (VLA), which was operated in a phased mode to act as a single element. Because of the wide bandwidth of the maser emission, two sets of interleaved observations were required; one to cover the systemic features and the red-shifted features, the other to cover the systemic features and the blue-shifted features. The data were processed on the new correlator of the VLBA, which generated 512 channel spectra for each of four 8-MHz bands resulting in a velocity resolution of 0.2 km s^{-1} . All spectra were phase referenced to the feature at 512.5 km s^{-1} to stabilize the interferometer. Synthesis images were constructed for all channels with the usual restoration techniques¹⁰. The synthesized beam was $0.6 \times 0.3 \text{ mas}$ at a position angle of $\sim 7^\circ$. We detect signal at $>5\sigma$ in 593 velocity channels (Fig. 1), out of 4,096 channels observed.

The spectroscopic images reveal the following physical characteristics (Table 1). The high-velocity features are offset from the systemic features in a nearly planar structure (Fig. 2a, b), and the velocity decreases with distance (r) from the centre of rotation as $r^{-1/2}$ (Fig. 3), exactly as expected for keplerian motion, with the features occurring in a line perpendicular to the line of sight. These features arise in a disk having inner and outer radii of ~ 4 and 8 mas , respectively. They define the systemic velocity (476 km s^{-1}), the binding mass ($3.6 \times 10^7 M_\odot$), the position angle of the disk (86°), and indicate that the disk is nearly edge-on. We estimated the inclination angle ($\sim 83^\circ$) from the displacement of the systemic features from the line between the high-velocity features. The systemic features, which are closer to us than the high-velocity features, show a systematic trend in velocity with position of $3.89 \text{ microarcsec per (km s}^{-1})$ along the major axis (Fig. 3 inset). The very tight linear correlation and the binding mass derived above indicate that all these features lie at a nearly constant radius of 4.1 mas (0.13 pc at our estimated

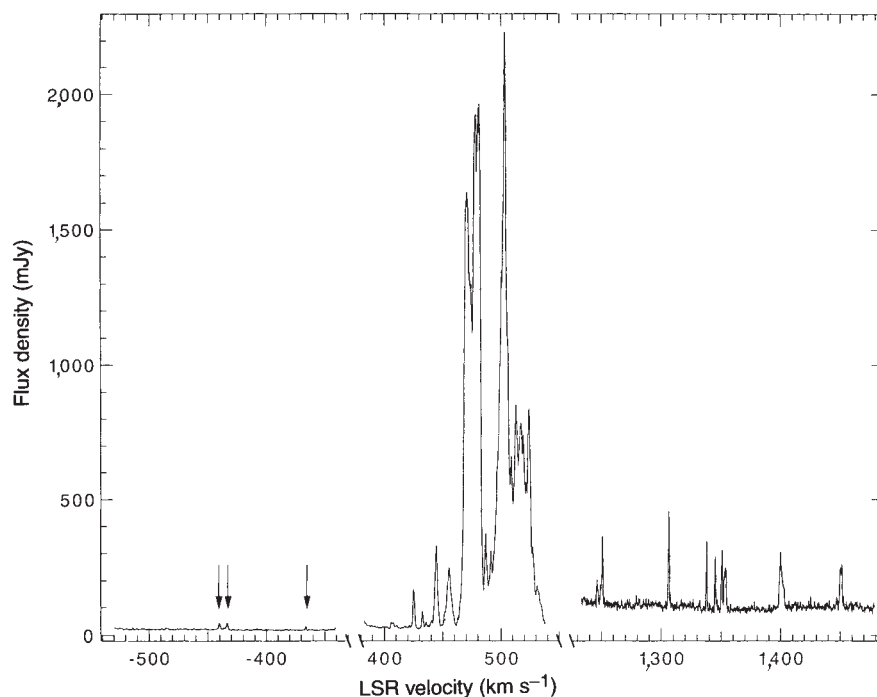


FIG. 1 Spectrum of the water-vapour maser in NGC4258 constructed from the interferometric images. The intensity is the maximum found in the $25 \times 25 \text{ mas}$ image at each velocity. The velocity scale is measured with respect to the local standard of rest (LSR). The rest frequency of the maser transition ($6_{16}-5_{23}$) is $22,235.08 \text{ MHz}$ (1.35 cm wavelength). The noise level is higher for the red-shifted features because the VLA did not contribute to that part of the spectrum. The biased baseline is an artefact of selecting the strongest feature in each velocity map.

distance of 6.4 Mpc), and the slope is caused by the change in projection of the velocity vector along the line of sight. The disk is very thin as the axial dimension is <0.1 mas so that height/radius ≤ 0.02 (see Fig. 2a, b), as expected for a disk in hydrostatic equilibrium¹¹.

Because the binding mass must be confined to a radius smaller than 4.1 mas, the resulting mass density is $>4 \times 10^9 M_\odot \text{pc}^{-3}$. It is unlikely that a star cluster could have such a high density (for example, globular clusters have stellar densities of $<10^5 M_\odot \text{pc}^{-3}$). If the mass were in a cluster of $1 M_\odot$ stars, the mean stellar separation would be only 100 AU, the collision time would be $<10^8$ yr; such collisions would disrupt the cluster. Hence the central mass is probably a black hole. Note that other black-hole candidates have substantially lower limits on their mass density (for example, M87^{12,13}, central mass $10^9 M_\odot$, central density $10^6 M_\odot \text{pc}^{-3}$; the Galactic Centre¹², central mass $10^6 M_\odot$, central density $10^8 M_\odot \text{pc}^{-3}$; NGC3115¹⁴, central mass $10^9 M_\odot$, central density $10^4 M_\odot \text{pc}^{-3}$). The mass of the central object is much greater than the mass of the disk. The disk does not cause a measurable deviation in the keplerian dependence of the high-velocity features, indicating that its mass is $<4 \times 10^6 M_\odot$. If we assume a disk thickness of 0.1 mas (0.003 pc) and a density of the disk of 10^{10} hydrogen molecules per cm^3 , the maximum likely value to avoid thermal quenching of the maser (removing the inversion in population levels and stopping the maser), the mass is $2 \times 10^5 M_\odot$.

Any radial motion (infall or outflow) is probably small compared to the rotational motion. The systematic velocity of the disk is very close to the systemic velocity of the galaxy ($472 \pm 4 \text{ km s}^{-1}$) (ref. 16), and to the systemic velocity estimated from the high-velocity features alone, which means that any radial motion in the systemic group is $\leq 10 \text{ km s}^{-1}$. The observed line-of-sight velocities of $\sim 1,000 \text{ km s}^{-1}$ of the high-velocity features are closely described by keplerian motion, and are unlikely to be due to outflow.

The restricted distribution of maser emission (Fig. 2c) may be related to physical conditions in the disk. The velocity gradient is

zero along the line of sight to the centre of the disk and along the radial line (mid-line) through the disk perpendicular to the line of sight. Hence the high-velocity features appear where the path for amplification is naturally very long ($\sim 10^{16} \text{ cm}$ for a velocity change of 1 km s^{-1} , the typical maser linewidth). The systemic features are clustered at the inner radius of the disk within a 8° angle, as viewed from the centre of the disk. This suggests that they may amplify a background continuum source (as suggested for the masers in Arp220¹⁷ and NGC3079¹⁸) of angular size ~ 0.5 mas, which means that their beam angle would be 7° . However, we did not detect any compact continuum emission at the level of 2 mJy (at a resolution of 1 mas), even though there is 3 mJy of emission from the nuclear region at 2-cm wavelength on the scale of 100 mas (ref. 19). Hence the maximum maser gain must be at least 1,000. The striking absence of features in the ranges from -400 to 410 km s^{-1} and from 530 to $1,100 \text{ km s}^{-1}$ is probably caused by the lack of a background source to amplify in our direction and the unfavourably short gain lengths. The emission in this part of the disk is beamed radially in other directions. The clustering of the systemic-velocity masers on the inner edge of the disk is interesting because the high-velocity features are distributed from radii of 4 to 8 mas. This clustering might be due to a beaming effect, as masers in the systemic group at the outer edge will be beamed into only 4° (because they are twice as far from the central background source as those on the inner radius) and their radiation might miss the Earth. Given a beam angle of 8° , the probability that we would see the maser emission from this galaxy (if it had a random inclination) is $\sim 6\%$. This is approximately the detection rate of megamasers in active galaxies^{4,20}. The nature of the pumping mechanism is unclear; however Neufeld *et al.*²¹ suggest that the X-ray emission from the nucleus could lead to a layer of excited water molecules in a shielded region of the molecular envelope, as well as a population inversion of the 1.35-cm transition.

The alignment of the spin axis of the molecular disk is offset from the spin axis of the galaxy by 119° (counter-rotating). Such

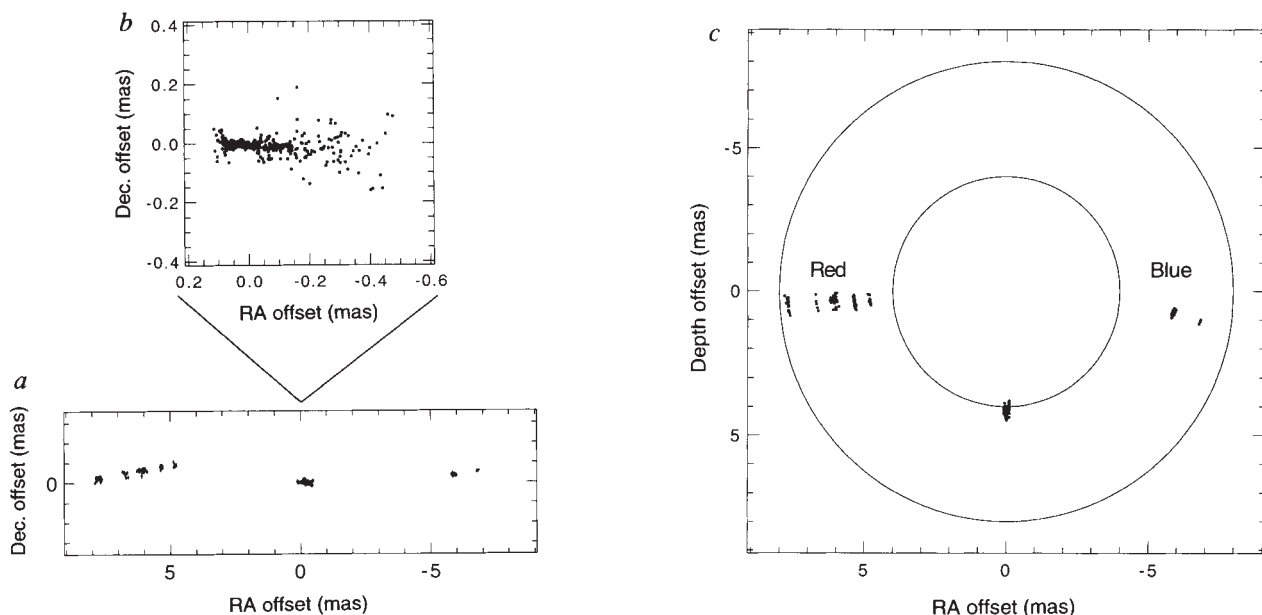


FIG. 2 a, Map of the distribution of the water masers in the NGC4258 (RA, right ascension; dec., declination). All features greater than 5σ are plotted (20–40 mJy). The position errors are $\sim 5 \times 10$ microarcsec for the systemic features, and $\sim 20 \times 40$ microarcsec for the high-velocity features. The origin of the coordinates is referred to the feature at 512.5 km s^{-1} . b, Expanded view of the emission near the systemic velocity of the galaxy. The features at $\text{RA} < -0.2$ mas are weak and the scatter in positions is due to measurement noise. c, Top view of the

molecular disk. The positions along the line-of-sight (depth) direction were calculated from the line-of-sight velocities and the calculated velocity field of a simple edge-on disk based on the binding mass required by the high-velocity features. Only those features that were measured accurately enough to provide errors in the depth direction of <0.15 mas are plotted. The circles mark the inner and outer edges of the observed molecular disk.

a misalignment is not uncommon between galaxies and their cores. But the spin axis of the disk is parallel, at least in projection, to the H α and radio synchrotron jets²² emanating from the nucleus on the scale of 500 pc. The molecular disk and helical strands have the same sense of rotation. We point out that there are systematic deviations from the simple disk model presented here, such as the departure from a linear dependence of the systematic features below 460 km s⁻¹ (Fig. 3 inset) and the deviations in the declinations of the high-velocity features from planar geometry (Fig. 2). The former effect is probably due to a slight fluctuation in the radii of the masers in that part of the disk (Fig. 2c). The latter effect is probably due to a warp^{23,24} or other distortion. It can be modelled as a series of concentric rings with constant inclination but with a progression in the line of nodes of $\sim 6^\circ$. None of the parameters in Table 1 are significantly affected.

Spectroscopic studies^{6,7} over several years show that the systemic features drift at a rate of 9.5 ± 1.1 km s⁻¹ yr⁻¹. Given the inner rotational velocity of 1,080 km s⁻¹, the velocity drift is undoubtedly due to the centripetal acceleration of gravity, $\dot{V} = V^2/R$ (where V is the rotational velocity at the inner radius R), from which we infer a linear radius of 0.13 pc. Combined with our angular size of 4.1 mas, we estimate the distance to be 6.4 ± 0.9 Mpc. Previous estimates have ranged from 3.3 to 7 Mpc (ref. 25). Our distance estimate is based on a direct geometric method, that is independent of the usual hierarchy of distance estimators. Its further refinement may help to improve the cosmic distance scale.

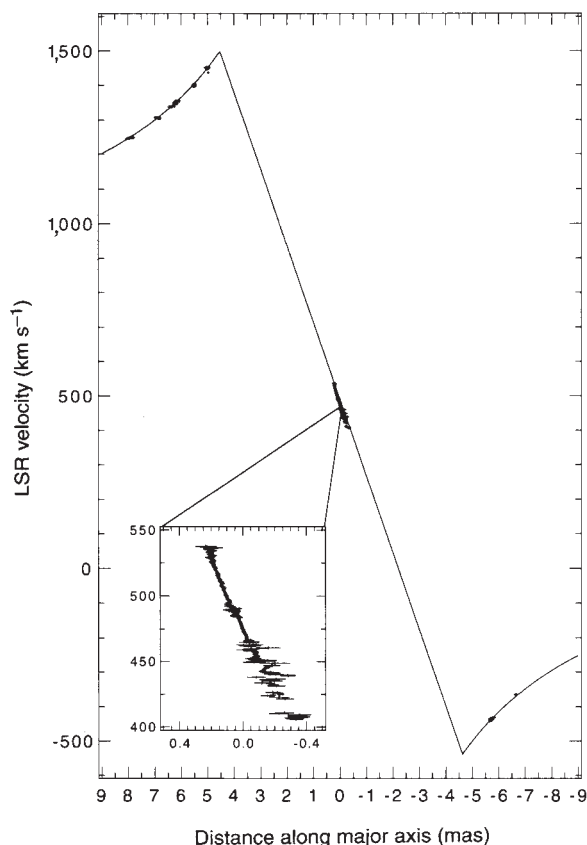


FIG. 3 Line-of-sight velocity versus distance along the major axis (position angle, 86°). Inset, data near the systemic velocity of the galaxy. The position errors are only visible on the scale of the inset. The line is derived from a model whose parameters are listed in Table 1. The high-velocity emission regions define keplerian orbits and constrain the estimate of the enclosed mass. The linear dependence of the systemic emission is a consequence of the change in projection of the rotation velocity.

TABLE 1 Parameters of molecular disk traced by water-vapour masers

Inner radius	4.1 mas (0.13 pc)*
Outer radius	8.0 mas (0.25 pc)*
Inner rotation velocity	$1,080 \pm 2$ km s ⁻¹
Outer rotation velocity	770 ± 2 km s ⁻¹
Inner rotation period	750 yr
Outer rotation period	2,100 yr
Position angle	$86 \pm 2^\circ$
Inclination	$83 \pm 4^\circ$
Position velocity slope (dr/dv)	3.89 ± 0.01 microarcsec per (km s ⁻¹)
Central mass	$3.6 \times 10^7 M_\odot$
Disk mass	$< 4 \times 10^6 M_\odot$
Central mass density	$> 4 \times 10^9 M_\odot \text{ pc}^{-3}$
Centripetal acceleration†	9.5 ± 1.1 km s ⁻¹ yr ⁻¹
Disk systemic velocity‡	476 ± 2 km s ⁻¹
Galactic systemic velocity‡§	472 ± 4 km s ⁻¹
Radial motion	< 10 km s ⁻¹
Thickness of disk	< 0.003 pc
Maser beam angle	7°
Disk-galaxy angle	119°
Apparent maser luminosity	$110 L_\odot$
Model luminosity¶	$7 L_\odot$
Distance	6.4 ± 0.9 Mpc

* Based on the distance estimate of 6.4 Mpc.

† From Greenhill *et al.*⁸.

‡ Radio definition, and with respect to the local standard of rest. To convert to heliocentric velocity (radio), subtract 8.2 km s⁻¹; to convert to heliocentric (optical), subtract 7.5 km s⁻¹.

§ From Cecil *et al.*¹⁶.

|| Angle between the spin axis of the molecular disk and the spin axis of the galaxy^{16,26}.

¶ Radiation into a zone within 4° of the plane of the disk.

The high-velocity features have velocity drifts of < 1 km s⁻¹ (refs 6, 7) as expected if they lie on the mid-line of the disk. They should also have small proper motions. The systemic features, however, should show proper motions of 35 microarcsec yr⁻¹. These could be readily measured in future VLBI observations in order to improve the definition of the disk and to estimate the distance more accurately. □

Received 2 December; accepted 21 December 1994.

- Rees, M. J. *A. Rev. Astr. Astrophys.* **22**, 471–506 (1989).
- Begelman, M. C., Blandford, R. D. & Rees, M. J. *Rev. mod. Phys.* **56**, 255–351 (1984).
- Claussen, M. J., Hellingman, G. M. & Lo, K. Y. *Nature* **310**, 298–300 (1984).
- Claussen, M. J. & Lo, K. Y. *Astrophys. J.* **308**, 592–599 (1986).
- Nakai, N., Inoue, M. & Miyoshi, M. *Nature* **361**, 45–47 (1993).
- Greenhill, L. J. *et al. Astrophys. J.* (in the press).
- Haschick, A. D., Baan, W. A. & Peng, E. W. *Astrophys. J.* **437**, L35–L38 (1994).
- Greenhill, L. J., Henkel, C., Becker, R., Wilson, T. L. & Wouterloot, J. G. A. *Astr. Astrophys.* (in the press).
- Watson, W. D. & Wallin, B. K. *Astrophys. J.* **432**, L35–L38 (1994).
- Zensus, J. A., Diamond, P. J. & Napier, P. J. *VLBI Summer School* (Astr. Soc. Pacific, San Francisco, in the press).
- Frank, J., King, A. & Raine, D. *Accretion Power in Astrophysics* 74 (Cambridge Univ. Press, 1992).
- Ford, H. C. *et al. Astrophys. J.* **435**, L27–L30 (1994).
- Harmes, R. J. *et al. Astrophys. J.* **435**, L35–L38 (1994).
- Kormendy, J. & Richstone, D. *Astrophys. J.* **393**, 559–578 (1992).
- Eckart, A., Genzel, R., Hofman, R., Sams, B. J. & Tacconi-Garman, L. E. *Astrophys. J.* **407**, L77–L80 (1993).
- Cecil, G., Wilson, A. S. & Tully, R. B. *Astrophys. J.* **390**, 365–377 (1992).
- Lonsdale, C. J., Diamond, P. J., Smith, H. E. & Lonsdale, C. J. *Nature* **370**, 117–120 (1994).
- Haschick, A. D. *et al. Astrophys. J.* **356**, 149–155 (1990).
- Turner, J. L. & Ho, P. T. P. *Astrophys. J.* **421**, 122–139 (1994).
- Henkel, C. *et al. Astr. Astrophys.* **141**, L1–L3 (1984).
- Neufeld, D. A., Maloney, P. R. & Conger, S. *Astrophys. J.* **436**, L127–L130 (1994).
- Cecil, G., Wilson, A. S. & DePree, C. *Astrophys. J.* (in the press).
- Binney, J. A. *Rev. Astr. Astrophys.* **30**, 51–74 (1992).
- Arnaboldi, M. & Galletta, G. *Astr. Astrophys.* **268**, 411–417 (1993).
- Courtès, G. *et al. Astr. Astrophys.* **268**, 419–442 (1993).
- Hubble, E. *Astrophys. J.* **97**, 112–118 (1943).

ACKNOWLEDGEMENTS. We thank G. Cecil, K. Kellermann, J. Kormendy, T. Nakano, M. Reid, R. Schild and B. Schmidt for helpful discussions. M. Miyoshi thanks the Foundation for the Promotion of Astronomy of Japan and the Smithsonian Institution for travel support. The VLBA and VLA are facilities of the National Radio Astronomy Observatory, which is operated by Associated Universities, Inc., under a cooperative agreement with the US NSF.

Friction melt distribution in a multi-ring impact basin

John G. Spray & Lucy M. Thompson

Department of Geology, University of New Brunswick, Fredericton, New Brunswick E3B 5A3, Canada

It is generally accepted that multi-ring basins are the consequence of very large impacts, but the mechanism by which they form is still a matter of contention^{1,2}. Most of what is currently known about multi-ring basins is based on remote studies of the Moon^{3–5} and, to a lesser extent, Mars and Mercury^{4,6}. But at least two multi-ring impact basins have been recognized on Earth⁷—the Sudbury⁸ (Canada) and Vredefort⁹ (South Africa) impact structures—providing an opportunity to study their properties directly. Here we describe the distribution of friction melt (pseudotachylite) in the floor of the Sudbury impact basin. Although the veins and dykes of pseudotachylite decrease in both thickness and frequency of occurrence towards the basin periphery, the greatest volumes of friction melt appear to define four rings around the central impact melt sheet. Field evidence indicates that the rings originated as zones of large displacement, which facilitated localized frictional melting of the basin floor during the modification (collapse) stage of the cratering process. By analogy, we argue that the rings of other multi-ring impact basins are also likely to be the remnants of such large-displacement fault zones.

The Sudbury impact structure, formed 1.85 Gyr ago, occurs in Archaean and Proterozoic lithologies of the Superior and Southern provinces of the Canadian Shield, northwest of the Grenville front (Fig. 1a). The core of the structure comprises a 60 × 30 km elliptical, 2.5 km thick, layered norite/quartz-gabbro/granophyre sequence known as the Sudbury igneous complex (SIC), interpreted to be an impact melt sheet^{8,10}. This is overlain by allochthonous breccia (suevite) and turbiditic sedimentary rocks of the Whitewater group. Post-impact deformation of the structure during the Penokean (spanning 1.9–1.6 Gyr ago) and, to a lesser extent, Grenvillian (~1.0 Gyr) orogenies produced the elliptical form of the SIC and caused north–south shortening and partial obliteration of the southern half of the impact structure^{11–13} (Fig. 1a). The SIC is surrounded by a number of ultra-high-strain-rate features indicative of shock and shock-related processes: multiple planar deformation features in quartz up to 8–10 km from the SIC contact with the basin floor¹⁴, shatter cones¹⁵ up to 20 km from the SIC and spectacular occurrences of pseudotachylite, locally referred to as Sudbury breccia, which pervades the basin floor^{16–18}. Our investigations have focused on determining the distribution of pseudotachylite north of the SIC, beyond the main overprinting effects of the Penokean and Grenvillian orogenies (Fig. 1a).

Pseudotachylite was first described and named by Shand¹⁹ based on his work on the Vredefort structure, where it occurs throughout the region as thin (<1 mm) veins and as larger dyke-like bodies several metres in thickness. The Sudbury pseudotachylites are similar to those found at Vredefort. They range in size from microscopic veinlets to large dykes several hundred metres in thickness and tens of kilometres in length. Most occurrences possess dark, fine-grained matrices with inclusions of adjacent wall rock and exotic material ranging from millimetre-sized mineral and rock fragments to metre-sized blocks¹⁸. It is possible to equate pseudotachylite development with the three main stages of impact cratering: (1) contact and compression, the initial period of projectile–target interaction; (2) excavation, with the production of a transient cavity which is gravitationally unstable and subsequently collapses during (3) the modification stage. Stages (1) and (2) are completed in seconds, whereas stage (3) may take several minutes or longer. Some of the submillimetre- to millimetre-thick pseudotachylites at Sudbury define

shatter-cone surfaces in the form of arcuate, overlapping veins, as has also been observed at Vredefort²⁰. Many of these smaller pseudotachylites do not cluster in or near fracture/slip systems associated with a given fault zone, but are instead pervasive on a massive scale as anastomosing networks. These may have been generated directly by the passage of the shock front during stages (1) and (2) and, if so, correspond to the type A pseudotachylites of Lambert²¹. These are particularly prevalent within 10–13 km of the SIC. The larger (mega-) pseudotachylite bodies, some of which are 400–500 m wide and up to at least 11 km in length (for example, the Froid–Stobie body west of the city of Sudbury), are typically associated with discrete planar surfaces, suggesting that they were generated by frictional heating along the fracture/slip planes that they now occupy. The larger pseudotachylites generally post-date the smaller type A veins and correspond to Lambert's type B variant associated with the modification stage of crater development²¹.

The majority of the Sudbury pseudotachylites possess matrices with micro-igneous textures, indicating that they cooled relatively rapidly from the molten state. The larger bodies are typically phaneritic and even medium-grained. The compositions of the Sudbury pseudotachylites show that they would have possessed very low viscosities when molten²². This is because the preferential melting of hydrous ferromagnesian phases in the wall rocks during pseudotachylite formation results in the generation of a relatively basic melt matrix, leaving predominantly quartz and feldspar as clasts²³. Furthermore, as clast–melt suspensions, pseudotachylites can undergo shear thinning during slip, resulting in an order of magnitude or more reduction in their viscosity²². The very low viscosities of the Sudbury pseudotachylites means that, once generated, they could have acted as effective fault lubricants and expedited large displacements. The more peripheral pseudotachylites tend to exhibit cataclastic textures and may have originated as gas-fluidized wall-rock powders. During high-speed slip (that is, >1 m s^{–1}), it has been demonstrated experimentally that comminution precedes frictional melting under conditions of continued or increasing energy dissipation at a sliding rock interface²⁴. It is not surprising, therefore, to find more evidence for cataclasis towards the edge of the impact basin where strain rates, although still very high, would have been lower than those realized toward the centre of the structure.

In terms of distribution, the amount of pseudotachylite in the basin floor decreases radially from the SIC. At distances >50 km its occurrence is sporadic. Detailed mapping along an 80-km radial swath in the North Range reveals the existence of four zones comprising larger pseudotachylite bodies and greater overall volumes of pseudotachylite veinlets pervading the basin floor (Fig. 1b). These occur at 0–13, 25–35, 42–48 and 78–80 km beyond the SIC. Assuming an undeformed radius of 30 km for the SIC, this constrains the original diameter of the Sudbury structure to ~225 km. The maximum individual pseudotachylite thickness developed within each of the zones decreases outwards, associated with a progressive decrease in overall zone thickness (that is, 13, 10, 8 and 2 km, respectively). Within the innermost zone (0–13 km wide) adjacent to the SIC, mesoscopic pseudotachylite veins are pervasive and it is difficult to stand on any outcrop without observing pseudotachylite at one scale or another. This inner zone is well developed around most of the SIC and we tentatively equate it with the outer part of the collapsed transient cavity which is otherwise obscured by the SIC. Significantly, if other confirmed pseudotachylite occurrences are combined with our radial data then a discontinuous subconcentric pattern is revealed at the regional scale (Fig. 1a). For example, the 25–35 km ring embraces pseudotachylite concentrations at the East Bull Lake and Shakespeare–Dunlop intrusions located west-southwest of the SIC²⁵, at Espanola southwest of the SIC, east of Highway 144 in the Emo, Rhodes and Botha townships²⁶, and ~20 km east of Lake Wanapitei²⁷. The outermost ring includes the pseudotachylite-rich zone located 80 km

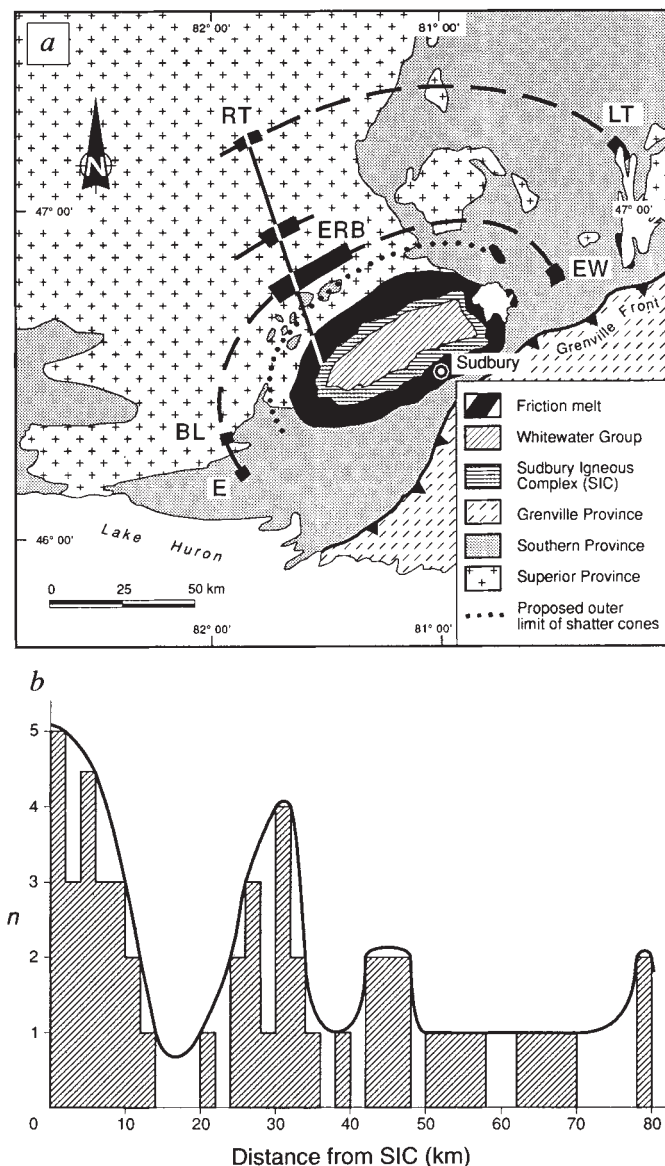


FIG. 1 *a*, Simplified geological setting of the 1.85 Gyr Sudbury igneous complex (SIC) and overlying Whitewater group within the Archaean Superior Province and Lower Proterozoic Southern Province of the Canadian Shield. Friction melt (pseudotachylyte) concentrations are shown based on detailed mapping along a radial transect (RT) in the North Range (Highway 144 and adjacent roads) and other known major occurrences at Espanola (E), East Bull Lake-Shakespeare Dunlop (BL), mbo-Rhodes-Botha townships (ERB), east of Lake Wanapitei (EW) and Lake Temagami (LT). *b*, Number *n* of pseudotachylytes (friction melts) encountered per 2 km plotted against distance from SIC along radial transect (RT) shown in *a*. Data based on counts for pseudotachylytes >1 cm thick. Concentrations occur at 0–13, 25–35, 42–48 and 78–80 km from the SIC.

northeast of the SIC in the Lake Temagami region²⁷. The dips of the pseudotachylyte-bearing faults are difficult to determine, though field and remote sensing evidence indicates that some are relatively steep²⁸. In terms of accommodating the collapse of the transient cavity and overall gravity-driven crater modification, we would expect the ring structures to assume progressively lower dip angles with depth in listric fashion. Ring-related scarp development has been largely obliterated in this eroded impact structure, but significant down-faulting is indicated south of the 25–35 km ring, where younger Huronian metasedimentary rocks have been preserved. Notwithstanding a lack of exposure, we acknowledge that the Sudbury rings are probably neither

continuous nor perfectly circular, but this is normal even for the type multi-ring basins of the Moon²⁹.

Previous workers studying lunar multi-ring basins have proposed that the rings follow a $\sqrt{2}$ distribution^{1,6}, although different spacings have also been proposed^{29,30}. Our ground observations reveal that the rings of the Sudbury structure do approximate a $\sqrt{2}$ pattern. This has been independently corroborated, at least for the inner two rings, by a recent Landsat-based lineament analysis of the Sudbury region made by Butler³¹.

Our evidence indicates that the majority of the Sudbury pseudotachylytes, and especially the larger dyke-like bodies, are the products of slip-induced frictional melting and are fault-generated. We see no reason to distinguish them from endogenic examples, with the exception of the type A pseudotachylyte variant which may be directly shock-generated and hence exclusive to impact craters. The main difference is one of scale: endogenic seismic faulting typically produces pseudotachylytes a few centimetres and locally up to a metre in thickness, whereas impact-induced pseudotachylytes can be up to hundreds of metres thick. For example, the amount of energy *E* required to produce a 100-m-thick friction melt is the product of the shear stress τ and total displacement *d*. For a coefficient of friction μ of 0.8, a density ρ of 3,000 kg m⁻³ and an estimated original burial depth *z* of 5 km, then $\tau \approx \mu \rho g z = 0.8 \times 3,000 \times 10 \times 5,000 = 1.2 \times 10^8$ Pa (where *g*, the acceleration due to gravity, has been approximated as 10 m s⁻¹). The energy required to melt 100 m of rock with a latent heat of fusion of 4×10^5 J kg⁻¹ = $100 \times 3,000 \times 4 \times 10^5 = 1.2 \times 10^{11}$ J m⁻². The amount of displacement required to produce a 100-m-thick friction melt during a single slip event is therefore $d = E/\tau = 1.2 \times 10^{11} / 1.2 \times 10^8 = 1,000$ m. This result is entirely compatible with observed scarp offsets of 6 km for lunar multi-ring impact basins²⁹, but is unreasonable for tectonic (endogenic) fault systems where single event displacements rarely exceed a few metres and so produce pseudotachylytes <1 m thick.

A number of theories have been proposed to explain the formation of multi-ring basins²⁹. One of the earliest was the volcanic modification theory, based on lunar observations³², whereby a crater is flooded with impact-triggered basaltic melt. Magmatic loading then induces subsidence via ring fractures (similar to terrestrial caldera collapse). This theory would hold only if the rings were formed after impact-triggered melt formation, yet the converse is true for Sudbury³³. The nested-crater hypothesis³⁴ explains the rings as crustal layer boundaries, which are exhumed by rebound, transient cavity collapse and subsequent isostatic uplift. At Sudbury, there is no evidence for the rings delineating crustal layering within the Archaean and Proterozoic target rocks. A popular idea has been that rings represent a "frozen tsunami wave" and that crater floors can behave as Bingham liquids on impact³⁵. The theory of acoustic fluidization explains this by proposing that impact-generated rock debris can flow as a bulk, low-viscosity fluid when subject to strong sound vibrations³⁶. Behaviour of the whole basin floor as a fluid is not borne out by our field and microscopic observations: apart from the presence of the pseudotachylytes, the target rocks retain normal igneous and metamorphic textures. Dence *et al.*³⁷, although not specifically addressing ring formation, suggest that crushing and frictional heating between basement blocks, with resultant frictional melting and pseudotachylyte injection, is a means of accommodating impact-induced displacement. We believe this is the most viable explanation of pseudotachylyte formation and, critically, our field observations link the focusing of the frictional melting process to the development of discrete, subconcentric zones of enhanced displacement.

Are there other examples of multi-ring basins on Earth? Vredefort is a likely contender, with rings located at 10–20 km and ~75 km, with at least one more at ~140 km. The more important implications of our Sudbury evidence are relevant to all the terrestrial planets. Previous definitions of ringed basins

have relied purely on their morphometry. This is unsatisfactory for the more eroded or obscured impact basins that are prevalent on tectonically active planets. Based on present definitions, these structures can never be classified because they no longer possess their original impact-generated topographies. New definitions should accommodate the physical character of the rings themselves and incorporate descriptions of the internal structure of impact basins. □

Received 27 June; accepted 25 November 1994.

- Hartmann, W. K. & Wood, C. A. *The Moon* **3**, 3–78 (1971).
- Spudis, P. D. *The Geology of Multi-Ring Impact Basins* 263 (Cambridge Univ. Press, New York, 1993).
- Hartmann, W. K. & Kuiper, G. P. *Comm. lunar planet. Lab.* **1**, 51–66 (1962).
- Pike, R. J. & Spudis, P. D. *Earth Moon Planets* **39**, 129–194 (1987).
- Baldwin, R. B. *The Measure of the Moon* 488 (Univ. Chicago Press, 1963).
- Wood, C. A. & Head, J. W. *Proc. lunar planet. Sci. Conf.* **7**, 3629–3651 (1976).
- Grieve, R. A. F., Robertson, P. B. & Dence, M. R. *Proc. lunar planet. Sci. Conf.* **12A**, 37–57 (1981).
- Grieve, R. A. F., Stöffler, D. & Deutsch, A. J. *geophys. Res.* **96**, 22753–22764 (1991).
- Martini, J. E. J. *metamorph. Geol.* **10**, 517–527 (1992).
- Faggart, B. E., Basu, A. R. & Tatsumoto, M. *Science* **230**, 436–439 (1985).
- Shanks, W. S. & Schwerdtner, W. M. *Can. J. Earth. Sci.* **28**, 1677–1686 (1991).
- Milkerit, B. et al. *Geology* **20**, 807–811 (1992).
- Thompson, L. M., Spray, J. G. & Kelley, S. P. (abstr.) *EOS* **74**, 389 (1993).
- Dence, M. R. *New Developments in Sudbury Geology* (ed. Guy-Bray, J. V.) 7–18 (Spec. Publ. No. 10, Geol. Ass. Can., Toronto, 1972).
- Dietz, R. S. *J. Geol.* **72**, 412–434 (1964).

- Speers, E. C. *J. Geol.* **65**, 497–505 (1957).
- Dressler, B. O. *The Geology and Ore Deposits of the Sudbury Structure* (eds Pye, E. G., Naldrett, A. J. & Giblin, P. E.) 97–136 (Ontario Geol. Surv. No. 1, Ont. Min. Nat. Res., Toronto, 1984).
- Thompson, L. M. & Spray, J. G. (Spec. Pap. 293, Geol. Soc. Am., in the press).
- Shand, S. J. *Q. Jl. geol. Soc. Lond.* **72**, 198–221 (1916).
- Martini, J. E. J. *Earth planet. Sci. Lett.* **103**, 285–300 (1991).
- Lambert, P. *Proc. lunar planet. Sci. Conf.* **12A**, 59–78 (1981).
- Spray, J. G. *J. geophys. Res.* **98**, 8053–8068 (1993).
- Spray, J. G. *Tectonophysics* **204**, 205–221 (1992).
- Spray, J. G. *J. struct. Geol.* **7**, 49–60 (1987).
- Chubb, P. T. et al. *Can. J. earth Sci.* (in the press).
- Dressler, B. O. *Geology of the Emo, Rhodes, and Botha Townships District of Sudbury 54* (Rep. No. 196, Ontario Geol. Surv., Ont. Min. Nat. Res., Toronto, 1980).
- Peredery, W. V. & Morrison, G. G. *The Geology and Ore Deposits of the Sudbury Structure* (eds Pye, E. G., Naldrett, A. J. & Giblin, P. E.) 491–511 (Ontario Geol. Surv. No. 1, Ont. Min. Nat. Res., Toronto, 1984).
- Dressler, B. O. et al. *Research in Terrestrial Impact Structures* (ed Pohl, J.) 39–68 (Vieweg, Braunschweig, Germany, 1987).
- Melosh, H. J. *Impact Cratering: A Geologic Process* 245 (Oxford Univ. Press, New York, 1989).
- Melosh, H. J. & McKinnon, W. B. *Geophys. Res. Lett.* **5**, 985–988 (1978).
- Butler, H. R. *Large Meteorite Impacts and Planetary Evolution* (eds Dressler, B. O., Grieve, R. A. F. & Sharpton, V. L.) (Spec. Pap. No. 293, Geol. Soc. Am., Boulder, CO, in the press).
- Hartmann, W. K. & Yale, F. G. *Sky Telesc.* **37**, 4–7 (1969).
- Wichman, R. W. & Schultz, P. H. *Meteoritics* **28**, 222–231 (1993).
- Hodges, C. A. & Wilhelms, D. E. *Icarus* **34**, 294–323 (1978).
- Baldwin, R. B. *Phys. Earth planet. Int.* **6**, 327–339 (1972).
- Melosh, H. J. *J. geophys. Res.* **84**, 7513–7520 (1979).
- Dence, M. R., Grieve, R. A. F. & Robertson, P. *Impact and Explosion Cratering* (eds Roddy, D. J., Pepin, R. O. & Merrill, R. B.) 247–275 (Pergamon, New York, 1977).

ACKNOWLEDGEMENTS. We thank J. Melosh, B. French and R. Grieve for reviews. This work was supported by LITHOPROBE (Abitibi-Grenville Transect), the Canadian NSERC and the Barringer Crater Co.

Direct spectroscopic observation of quantum jumps of a single molecule

Th. Basché, S. Kummer & C. Bräuchle

Institut für Physikalische Chemie, Universität München,
Sophienstrasse 11, 80333 München, Germany

BOHR'S notion of quantum jumps between electronic states of an excited atom has now been demonstrated experimentally for single ions confined in radio-frequency traps and interacting with a driving laser field^{1–3}. In these experiments the fluorescence of a strongly allowed transition was shown to cease abruptly when the ion jumped into a metastable state which was coupled to the common electronic ground state by a weak radiative transition. But attempts to monitor quantum jumps of single molecules have been hampered by the fact that the lifetime of the metastable triplet state was too short in relation to the photon detection rate. By using a system with favourable photophysical parameters—terrylene doped into *p*-terphenyl crystals⁴—we have now been able to observe directly quantum jumps between electronic states of single terrylene molecules. In contrast to single atoms, here the quantum jumps occur as non-radiative transitions between states of different multiplicity, and are manifested as interruptions of the fluorescence signal. These results demonstrate how single-molecule spectroscopy can reveal truly quantum-mechanical effects in large polyatomic molecules.

Crystals of terrylene (C₃₀H₁₆, a higher homologue of perylene) in *p*-terphenyl were grown by sublimation from a mixture of the two compounds. The crystals were mounted in a cryostat (1.4 K) at the joint focus of a lens and a paraboloid which served as an efficient collector of the fluorescence emission. To detect single terrylene molecules, a single-frequency rhodamine 6G dye laser was scanned over the spectral region around 578.48 nm while simultaneously recording the Stokes-shifted terrylene fluorescence with single-photon-counting electronics after red-pass filtering. The terrylene concentration in our samples was so low that the single-molecule excitation lines appeared as isolated features distributed over ~50 GHz. A detailed description of the

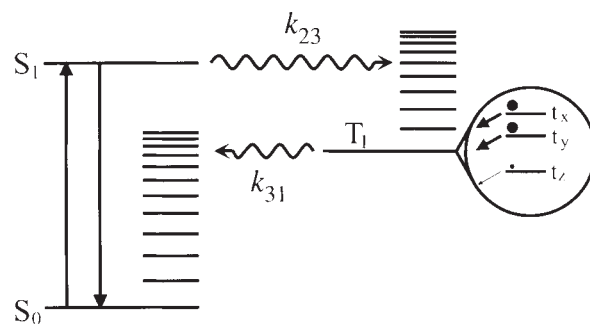


FIG. 1 Energy-level scheme for terrylene. S_0 denotes the electronic ground state, S_1 the first excited singlet state and T_1 the first excited triplet state. The molecule is irradiated with laser light in resonance with the $S_0 \rightarrow S_1$ zero-phonon transition. From S_1 it may fluoresce or decay to T_1 . The probability of the latter process is extremely small ($\sim 10^{-6}$) because intersystem crossing (ISC) from S_1 to T_1 is spin-forbidden. As the same is true for ISC from T_1 to S_0 , the lifetime of T_1 is orders of magnitude longer than that of S_1 . ISC is a non-radiative process which is thought to occur between isoenergetic levels of the singlet and triplet manifold. As S_1 and T_1 are different in energy, vibrational quanta are always involved in ISC. The transition rates k_{23} and k_{31} determine the ISC dynamics because the vibrational relaxation of a large molecule in a solid is extremely fast (picoseconds). The triplet state T_1 is shown on an enlarged scale at the extreme right of the figure. It is split into three sublevels by magnetic dipolar interaction of the triplet electrons (zero-field splitting). The dots and arrows denote the approximate populations and lifetimes of the sublevels as derived from fluorescence correlation measurements on single terrylene molecules.

apparatus and the principle of single-molecule spectroscopy is given in refs 5 and 6.

The discrete turning on and off of the strong fluorescence of an allowed electronic transition by shelving of the optical electron in a long-lived metastable state is a true signature of quantum jumps of a single particle⁷. In experiments with single atoms, the metastable (or dark) state was populated by illuminating the atom with a second light field, in addition to the driving laser field which populated the highly emissive state. The situation is different for organic dye molecules with an electronic level structure as described in Fig. 1. In such a three-level scheme, the metastable triplet state serves as the dark state which is populated and depopulated by non-radiative transitions⁸.

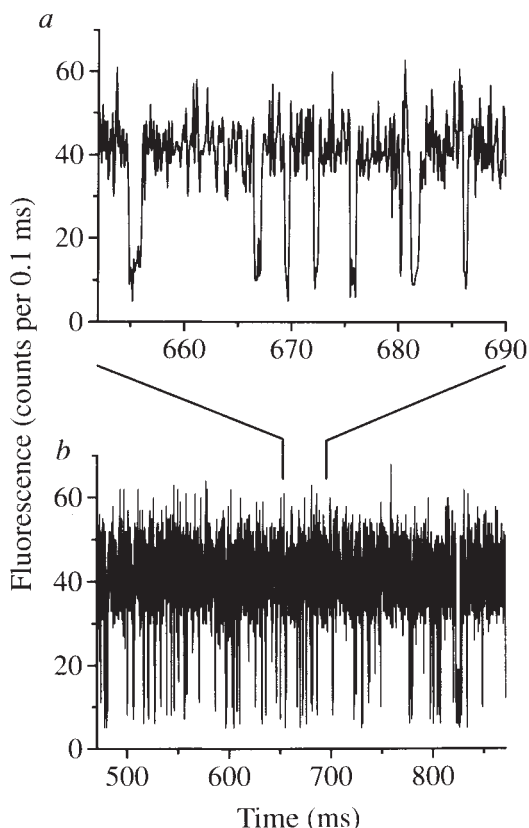


FIG. 2 Fluorescence counts detected for a single terrylene molecule as a function of time. The quantum jumps are clearly visible as discrete interruptions of the fluorescence. The fluorescence counts were recorded with a multichannel scaler set to a sampling interval of 0.1 ms; this instrument did not introduce any delay between successive sampling intervals. *b*, a 400-ms section of a 1.64-s data set; *a*, a region covering 38 ms displayed on a magnified scale. The non-zero count rate when the molecule is shelved in the triplet state is caused by laser light scattered from the host crystal. The sample was cooled to 1.4 K and the molecule was excited at a laser wavelength of 578.494 nm with an intensity of 25 W cm^{-2} .

Quantum jumps between the singlet and triplet manifold of a single molecule should induce large fluctuations in the fluorescence intensity of the molecule, an effect which is also termed photon bunching. An indirect means to look at quantum transitions of a single molecule is to record the fluorescence intensity autocorrelation function⁹. Such measurements were done with single pentacene molecules in *p*-terphenyl, allowing the determination of the intersystem crossing (ISC) rates k_{23} and k_{31} from observations of photon bunching on the microsecond timescale⁹, and permitting the observation of the non-classical effect of photon antibunching¹⁰ on the nanosecond timescale (of the same order as the lifetime of the singlet excited state S_1).

As k_{23} and k_{31} of terrylene were not known from experiments on ensembles of molecules, we extracted them from correlation measurements on single terrylene molecules⁴. In Fig. 1 it is seen that the triplet state is split into three sublevels which, because of the symmetry selection rules of the ISC, have in general different population and depopulation rates¹¹. We found that the kinetics for the in-plane levels t_v and $t_{v'}$ are indistinguishable in our experiments and a factor of ten faster than the kinetics of t_z ; this is typical of aromatic hydrocarbons¹². For the analysis of the experiments presented below it is sufficient to consider the fast rates only because they dominate the ISC. Typical values obtained were $k_{23} \leq 10^3 \text{ s}^{-1}$ and $k_{31} \approx 2 \times 10^3 \text{ s}^{-1}$. This means that we expect a molecule to undergo excitation–emission cycles in the singlet system for (on average) $\sim 2 \text{ ms}$ ($2/k_{23}$), and after ISC into the triplet state it will be shelved for $\sim 0.5 \text{ ms}$ ($1/k_{31}$).

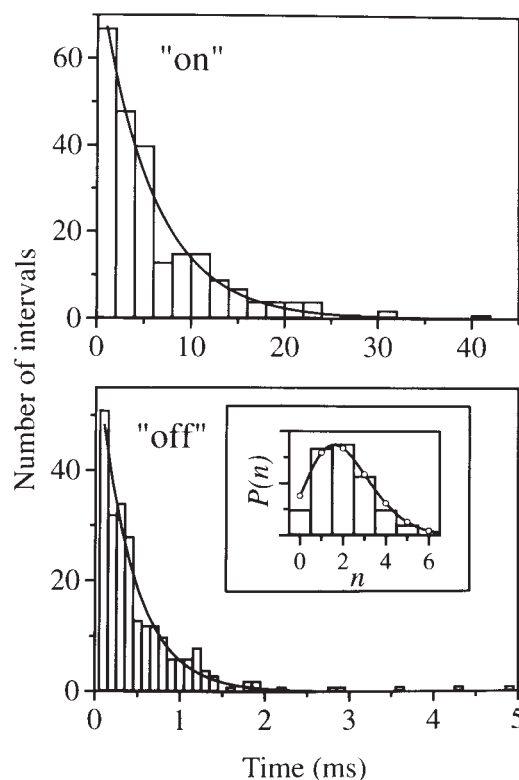


FIG. 3 Distributions of the lengths of 'on' and 'off' intervals (top and bottom panels, respectively). The lines are exponential fits to the histograms. Inset, the distribution of the number of jumps n out of the dark state in 15 ms time increments. The line is the result of fitting a poissonian distribution $P(n)$ to the data. The source of the data shown here was a trace of 1.64 s length, part of which is shown in Fig. 2.

Considering the above parameters and a typical number of detected fluorescence photons of $\sim 5 \times 10^5$ per second for a single terrylene molecule, any abrupt fluctuations of the fluorescence should be easily observable when sampling the fluorescence intensity within time intervals of 50–100 μs .

Figure 2 shows the average fluorescence intensity of a single terrylene molecule as a function of time; the abrupt macroscopic interruptions of the single-molecule fluorescence are clearly visible (Fig. 2*a*). The single molecule displays individual quantum jumps by turning on and off the strong fluorescence of the S_0 – S_1 transition. This random 'telegraph signal' tells us the quantum state of the molecule: whether it is in the singlet manifold or shelved in the triplet state. The precise knowledge of the quantum state of the molecule at any instant of the experiment is an important advantage of the direct quantum-jump measurement when compared to the correlation technique⁹. For instance, timing of optically detected magnetic-resonance experiments^{13,14} can be performed very accurately because the detection efficiency for every transition to the triplet state is unity provided that the residence time in T_1 is longer than the sampling time for the fluorescence.

We can rule out the possibility that the fluctuations of the fluorescence are caused by spectral diffusion of the molecular resonance line⁵. Such diffusion depends on the local surroundings of the molecule and its time dependence should strongly vary from molecule to molecule. The jump rates into and out of T_1 , however, were found to be very similar for all single terrylene molecules investigated.

Figure 3 shows the distributions of interval lengths, derived from a trace with recording time of 1.64 s, for the situations when the molecule is in the singlet ('on'-times) or in the triplet

manifold ('off' -times), respectively. These distributions should be exponential and allow the determination of the rates of the underlying processes¹⁵. From exponential fits to the distributions we obtained $\tau_{\text{off}} = 0.4$ ms and $\tau_{\text{on}} = 6.9$ ms. τ_{off} is independent of intensity and equals the lifetime of T_1 ($1/k_{31}$) whereas $\tau_{\text{on}} = 2/k_{23}$ only when the $S_0 \rightarrow S_1$ transition is fully saturated. When the intensity is increased by a factor of 2.5 we reach the saturated regime, and as expected τ_{off} remains constant while τ_{on} decreases to 4.5 ms. From a measurement of the correlation function of the same molecule we obtained $k_{23} = 0.4 \times 10^3 \text{ s}^{-1}$ and $k_{31} = 2.5 \times 10^3 \text{ s}^{-1}$ which is in excellent agreement with the data derived from the direct quantum-jump measurement. Finally, Fig. 3 inset shows the distribution of the number of jumps out of the dark state per finite time interval. This distribution obeys the expected poissonian statistics⁷ as can be seen from the corresponding fit.

The ability to directly observe quantum jumps between electronic levels with different multiplicity is inevitably connected to the internal photophysical dynamics of the molecular system. One can envisage constructing a single-atom (molecule) optical clock—an idea which is based on a three-level scheme¹ inherent to any dye molecule—if one can identify systems with much longer lifetimes of the metastable state. In such a system, using the macroscopic fluorescence signal one would monitor the presence of long-lived metastable states. The longer the lifetime, the narrower the linewidth, and in the limit of very narrow linewidths one might hope to develop a very precise frequency standard. For terrylene the average length of 'off' times could be influenced by applying microwaves which couple the triplet sub-

levels and therefore change the dynamics¹⁴. With accurate timing this would, however, lengthen the dark intervals only up to an order of magnitude considering the known kinetics. Nevertheless, with the proper choice of system, a single molecule or ion doped into a solid and cooled to low temperatures may be an interesting alternative to a single laser-cooled ion in a trap. As an additional advantage, the molecule would be essentially at rest and collision-induced jumps would be absent. □

Received 10 October; accepted 7 December 1994.

1. Nagourney, W., Sandberg, J. & Dehmelt, H. *Phys. Rev. Lett.* **56**, 2797–2799 (1986).
2. Sauter, Th., Neuhauser, W., Blatt, R. & Toschek, P. E. *Phys. Rev. Lett.* **57**, 1696–1698 (1986).
3. Bergquist, J. C., Hulet, R. G., Itano, W. M. & Wineland, D. J. *Phys. Rev. Lett.* **57**, 1699–1702 (1986).
4. Kummer, S., Basché, Th. & Bräuchle, C. *Chem. Phys. Lett.* **229**, 309–316 (1994).
5. Ambrose, W. P., Basché, Th. & Moerner, W. E. *J. chem. Phys.* **95**, 7150–7163 (1991).
6. Moerner, W. E. & Basché, Th. *Angew. chem.* **105**, 537–557 (1993); *Angew. Chem. int. Edn engl.* **32**, 457–476 (1993).
7. Schenzle, A. & Brewer, R. G. *Phys. Rev. A* **34**, 3127–3142 (1986).
8. Birks, J. B. *Photophysics of Aromatic Molecules* (Wiley-Interscience, London, 1970).
9. Bernard, J., Fleury, L., Talon, H. & Orrit, M. *J. chem. Phys.* **98**, 850–859 (1993).
10. Basché, Th., Moerner, W. E., Orrit, M. & Talon, H. *Phys. Rev. Lett.* **69**, 1516–1519 (1992).
11. McGlynn, S. P., Azumi, T. & Kinoshita, M. *Molecular Spectroscopy of the Triplet State* (Prentice Hall, Englewood Cliffs, 1969).
12. Lawetz, V., Orlandi, G. & Siebrand, W. J. *chem. Phys.* **56**, 4058–4072 (1972).
13. Köhler, J. et al. *Nature* **363**, 242–244 (1993).
14. Wrachtrup, J., von Borczkowski, C., Bernard, J., Orrit, M. & Brown, R. *Nature* **363**, 244–245 (1993).
15. Cook, R. J. & Kimble, H. J. *Phys. Rev. Lett.* **54**, 1023–1026 (1985).

ACKNOWLEDGEMENTS. Financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

Relation between century-scale Holocene arid intervals in tropical and temperate zones

H. F. Lamb*, F. Gasse†, A. Benkaddour*†, N. El Hamouti†, S. van der Kaars*, W. T. Perkins*, N. J. Pearce* & C. N. Roberts†

* Institute of Earth Studies, University of Wales, Aberystwyth SY23 3DB, UK

† Laboratoire d'Hydrologie et de Géochimie Isotopique, Centre Scientifique d'Orsay, Université de Paris-Sud, 91405 Cedex, France

‡ Department of Geography, Loughborough University of Technology, Loughborough LE11 3TU, UK

CLIMATE records from lake sediments in tropical Africa, Central America and west Asia show several century-scale arid intervals during the Holocene^{1–10}. These may have been caused by temporary weakening of the monsoonal circulation associated with reduced northward heat transport by the oceans⁷ or by feedback processes stimulated by changes in tropical land-surface conditions¹⁰. Here we use a lake-sediment record from the montane Mediterranean zone of Morocco to address the question of whether these events were also felt in temperate continental regions. We find evidence of arid intervals of similar duration, periodicity and possibly timing to those in the tropics. But our pollen data show that the forest vegetation was not substantially affected by these events, indicating that precipitation remained adequate during the summer growing season. Thus, the depletion of the groundwater aquifer that imprinted the dry events in the lake record must have resulted from reduced winter precipitation. We suggest that the occurrence of arid events during the summer in the tropics but during the winter at temperate latitudes can be rationalized if they

are both associated with cooler sea surface temperatures in the North Atlantic.

Tigalmamine (32° 54' N, 5° 21' W, area 6 ha, maximum depth 16 m) is the largest of a group of solution lakes in Lower Lias dolomite bedrock, at 1,628 m above sea level in the Middle Atlas of Morocco. The lake (Table 1) is fed by ground water, surface springs, and winter runoff from the 3.5 km² catchment; its surface outlet is dry in summer. The climate of the area is strongly seasonal (Table 1): 90% of the total mean annual precipitation of 930 mm falls between October and May, from westerly winds¹¹. Evergreen oak (*Quercus rotundifolia*) and Atlas Cedar (*Cedrus atlantica*) dominate the forest vegetation. Eleven radiocarbon ages (Table 1) on a 16-m-long core from the lake centre show that sediment deposition was continuous throughout the Holocene, in contrast to a littoral core^{12,13}. The endogenic origin of the dated fine-fraction (<80 µm) calcite is confirmed by scanning electron microscope observations, and by its $\delta^{13}\text{C}$ values, which are closer to $\delta^{13}\text{C}$ values of the present-day total dissolved inorganic carbon (TDIC) of the lake than to those of catchment rock and soil carbonate (Table 1). The calcite ^{13}C content is negative relative to calcite precipitated in equilibrium with atmospheric CO_2 (+1 to +4‰; ref. 14), so we used a mixing model^{15,16} to correct the radiocarbon dates for the hardwater effect of incomplete equilibration of lakewater TDIC with atmospheric CO_2 . The corrections (of 300–700 yr) are consistent with ages obtained for present-day lakewater TDIC. The chronology is based on these corrected calcite dates, because the provenance of dated organic material is less clear.

Most of the core (units 3, 5, 7, 9, 10, 11; Fig. 1) consists of dark-brown, organic, clay-rich marl with frequent millimetre-scale calcite laminations, similar to sediments forming at present. Calcite, as rhomboidal crystals of endogenic epilimnetic origin, is the dominant mineral constituent; aragonite, from molluscan shell fragments, is present in trace amounts. Clays, quartz and dolomite probably entered the lake by way of runoff during periods of high lake-level. In contrast, units 2, 4, 6 and 8, delimited by sharp lithologic boundaries, are composed of coarsely granular calcite, including abundant casts and oogonia of the

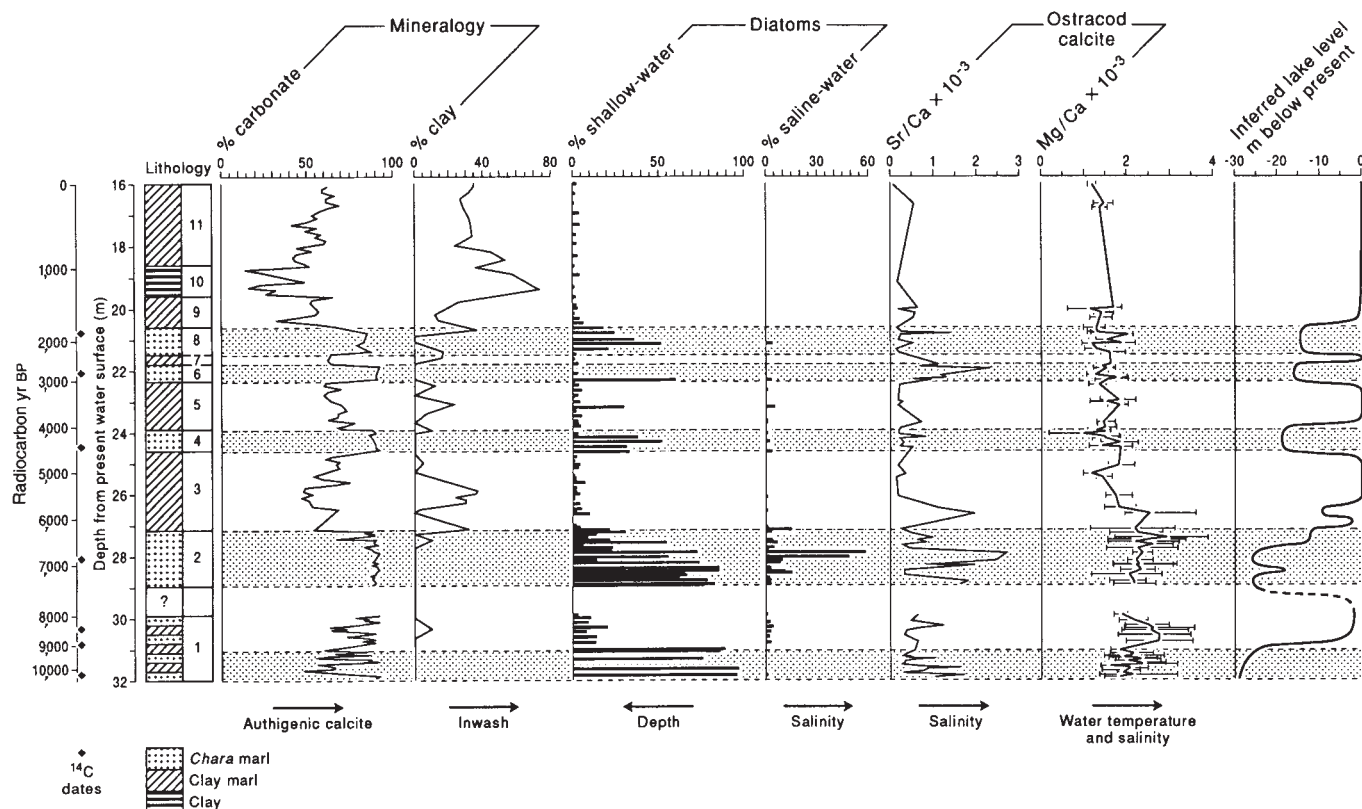


FIG. 1 Lithology, radiocarbon chronology and biostratigraphy of the Tigalmamine C86 core. Sr/Ca and Mg/Ca element concentration ratios, determined by inductively coupled plasma mass spectrometry, are

means of three measurements on individual valves; the ranges are shown for Mg/Ca. Inferred shallow-water phases are indicated by shading.

macrophytic alga *Chara*, ostracods (mostly *Fabaeformiscandona fabaeformis*, characteristically active all year in shallow ponds, with ecdysis between May and October^{17,18}), and gastropods. Detrital clays, quartz and dolomite are virtually absent from these *Chara*-marl units, which are interpreted as shallow-water facies, formed during low lake-level stands at times of reduced runoff. *Chara* grows only in shallow water,^{19,20} not below 8 m depth in the present lake. Unit 1 is a complex series of alternating *Chara*- and clay-marls, suggesting initially unstable water levels in the lake.

The lithostratigraphic interpretation is supported by the dominance of periphytic, shallow-water diatoms (*Cocconeis placentula*, *Gomphonema angustatum*, *Epithemia zebra*) in units 2, 4 and 8 (Fig. 1; diatom preservation is poor in unit 6; ref. 21). Unit 2 includes successive peaks of *Cyclotella* sp. aff. *antiqua* and *Mastogloia smithii*; the latter tolerates slightly saline conditions. The intervening deep-water facies are characterised by planktonic diatoms such as *C. parva* and *C. azigzensis*. The diatoms suggest that shallow-water conditions prevailed during deposition of the lowest metre of unit 1. We also determined Sr/Ca ratios on each of three valves of the ostracod *F. fabaeformis* at 80 levels in the core. These ratios show peak values during the inferred shallow-water intervals, especially those of units 2 and 6 (Fig. 1). Because the trace-element content of ostracod-shell calcite reflects water chemistry at the time of valve formation²²⁻²⁴, the peak Sr/Ca ratios suggest slight evaporative concentration of the lake water during low-level phases. The Mg/Ca ratio of ostracod calcite is related to both salinity and ambient water temperature²⁴, but because Ca concentrations are not constrained by carbonate saturation at low salinities, this ratio shows a weak response to the inferred shallow-water intervals. Nevertheless, its profile shows higher mean and greater range values until the base of unit 3, which may reflect a smaller

lakewater volume during the early Holocene, or possibly enhanced temperature seasonality relative to the present²⁵. The exclusion of *Cedrus atlantica*, which now occurs in cool and moist habitats between 1,500 and 2,100 m (ref. 26), until 6,500 ¹⁴C yr before present (BP) and its low abundance until 4,500 ¹⁴C yr BP (Fig. 2) may also be a result of higher Early Holocene summer temperatures.

We interpret these stratigraphic changes as a record of lake level and salinity responses to Holocene humid-arid climate variation, mediated by ground water and hydrological mass-balance. The record suggests that the level of Tigalmamine was substantially lower than present during five intervals lasting between 150 and 400 years (Table 2). The lake may not have recorded climate changes during humid phases, because it cannot have responded to precipitation/evaporation changes until its level fell below the overflow threshold. As the lake filled with sediment, smaller lake-level falls would have resulted in shallow-water conditions suitable for the growth of *Chara*, so the sediments became progressively more sensitive to changes in hydro-climatic balance. Thus, the earlier part of the sequence records only the most intense regressional episodes; later Holocene events have an apparently lower magnitude.

It seems that even prolonged drought had limited effect on the forest vegetation. Deciduous oak (*Quercus canariensis*), which is less drought-tolerant than the other tree taxa recorded, shows slightly decreased percentage abundances during the regressional episodes of units 4, 6, and 8 (Fig. 2). The other taxa apparently suffered no ill effects from aridity. The longevity (up to 1,500 yr) of individual *Cedrus* trees exceeds the duration of the arid intervals, so although drought may have delayed recruitment, it did not increase mortality. Precipitation may therefore have remained sufficient to sustain plant growth during the summer growing season. This suggests that reduced precipitation in

TABLE 1 Modern characteristics and ^{14}C -dated material from Tigalmamine

Mean air temperature (°C)		Mean monthly precipitation (mm)			
January	2.5	June–September	23		
July	20	October–May	126		
Lake property	Depth				
	Surface	8.5m(thermocline)	16 m		
pH	8.3	7.8	7.3		
Temperature (°C)	21.5	15.5	8.5		
Conductivity (μS cm ^{−1})	700	700	750		
Ca (mg l ^{−1})	68				
Mg (mg l ^{−1})	36.6				
Core C86; sample depth from water surface (cm)	Sample material analysis method	Conventional radiocarbon age (yr BP)	Hardwater-corrected radiocarbon age (yr BP) (refs 15, 16)	Calibrated age ³⁵ (1σ range) (cal yr BP)	δ ¹³ C (‰)
2,070–2,080	Organic AMS	2,990 ± 100		3,170 (3,330–2,990)	
2,070–2,080	Calcite, LCS	2,230 ± 140	1,755 ± 140	1,640 (1,840–1,520)	−7.07
2,200–2,210	Organic, AMS	3,220 ± 100		3,440 (3,550–3,350)	
2,200–2,210	Calcite, LCS	3,520 ± 100	2,840 ± 100	2,940 (3,080–2,190)	−7.06
2,438–2,448	Organic AMS	4,440 ± 100		5,020 (5,280–4,870)	
2,438–2,448	Calcite, LCS	4,800 ± 100	4,415 ± 100	4,980 (5,260–4,860)	−7.85
2,800–2,812	Calcite, LCS	7,300 ± 140	6,865 ± 140	7,650 (7,790–7,540)	−7.52
2,872	Organic, AMS	7,760 ± 110		8,490 (8,580–8,400)	
3,030	Calcite, AMS	8,510 ± 90	8,240 ± 90	9,220 (9,370–9,000)	−7.62
3,080	Calcite, AMS	9,150 ± 90	8,920 ± 90	9,940 (9,980–9,880)	−7.48
3,173–3,183	Calcite, AMS	10,390 ± 250	10,220 ± 250	12,004 (12,390–11,070)	
Lake surface water	TDIC AMS	650 ± 110			−9.80
Rock and soil carbonates					+0.33

Data taken from refs 15 and 21. LCS, liquid scintillation spectrometry; AMS, accelerator mass spectrometry, both carried out at the Laboratoire d'Hydrologie et de Géochimie Isotopique, Orsay. ($\delta^{13}\text{C} = [(^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{standard}}] - 1$; the PDB standard was used.)

winter caused the groundwater aquifer to become depleted; this would have lowered lake level without affecting the vegetation, because the mountain winters are too cold for plant growth.

Anthropogenic exploitation of the lake catchment is apparent from about 1,500 ^{14}C yr BP: the high clay content of unit 11 reflects soil erosion following disturbance. Deciduous oak declined more than the evergreen *Q. rotundifolia*, because the latter sprouts more readily after cutting and browsing. The accompanying Gramineae (grass) pollen increase records arable and pastoral agriculture, but an earlier grass pollen rise ($\sim 6,000$ ^{14}C yr BP) is more likely to be a reflection of the initial development of lake-marginal *Phragmites* marsh on the valley floor during the first high-water phase.

The regressional events represent short-term decreases in winter rainfall during at least five episodes of the past 10,000 years (Table 2) that may be related to sea surface conditions in the North Atlantic. The salt-oscillator model of thermohaline circulation²⁷ requires approximately the time interval between events observed in the Tigalmamine record. A marine record from the Rockall Plateau²⁸ documents low summer sea surface temperature (SST), low salinity events centred on 11–10.5, 8–7 and 4.5–3 ^{14}C kyr BP. However, a record closer to Morocco, from a core off Portugal^{28,29}, shows slightly different timing with similar events at 11.5–11, 8.5–8, and a low-magnitude shift at ~ 4 ^{14}C kyr BP. Winter SSTs inferred from that core do not show

significant changes over the past 9.4 ^{14}C kyr, which may be due to the relatively low resolution of the marine records, and/or radiocarbon chronological uncertainties on both marine and lacustrine records. Low lake-level events of a similar duration, periodicity and possibly timing (a careful assessment and more precise chronology is still required) have been recorded from the north Sahara³⁰ and from tropical lakes^{1–10}, where they have been attributed to temporary weakening of the monsoonal circulation in response to decreased efficiency of the thermohaline conveyor belt⁷, to volcanic activity³¹, or to changes in tropical land conditions and related feedback processes¹⁰, amongst other hypotheses⁹. Reduced oceanic heat transfer may have resulted from glacial meltwater influx to the North Atlantic inhibiting deep-water formation for the 11–10 (Younger Dryas) and 7.5 kyr events⁷, whereas another mechanism is required to explain the Late Holocene arid intervals.

The similarities between arid episodes in Morocco and in the tropics suggests that both tropical summer and middle-latitude winter events were responses to changes in Atlantic heat transport, despite the lack of significant correlation between recent rainfall variation in Morocco and other parts of Africa^{32,33}. Other high-resolution middle-latitude continental and marine records are required to advance understanding of the causes and ecosystem responses of these abrupt climate changes, which probably have more immediate importance in predicting the

TABLE 2 Ages and durations of arid intervals at Tigalmamine

Lithological unit	Depth (cm)	Apparent age (duration) (corrected ^{14}C yr BP)	Adjusted age (duration) (^{14}C yr BP)	Adjusted age (duration) (cal yr BP)
8	2,060–2,148	1,700–2,350 (650)	1,750–1,950 (200)	1,660–1,880 (220)
6	2,180–2,235	2,650–3,050 (400)	2,750–2,900 (150)	2,830–2,990 (160)
4	2,390–2,460	4,050–4,550 (500)	4,300–4,450 (150)	4,860–5,010 (150)
2	2,716–>2,897	6,250–>7,400 ($\geq 1,150$)	6,650–>7,050 (≥ 400)	7,500–>7,860 (≥ 360)
1(lower half)	3,100–3,200	9,200–10,550 (1,350)		10,230–>12,480 ($\geq 2,250$)

These ages and durations are adjusted on the basis that low pollen concentrations (Fig. 2) of the shallow-water units 2–8 indicate that they accumulated three times faster than the mean rate of 1.4 mm yr^{-1} . The low pollen concentrations are due to the diluting effect of enhanced sediment formation rather than to a genuine decrease in pollen production by the vegetation, which would not be similar amongst taxa of various drought sensitivities. The base of unit 2 is missing from the record.

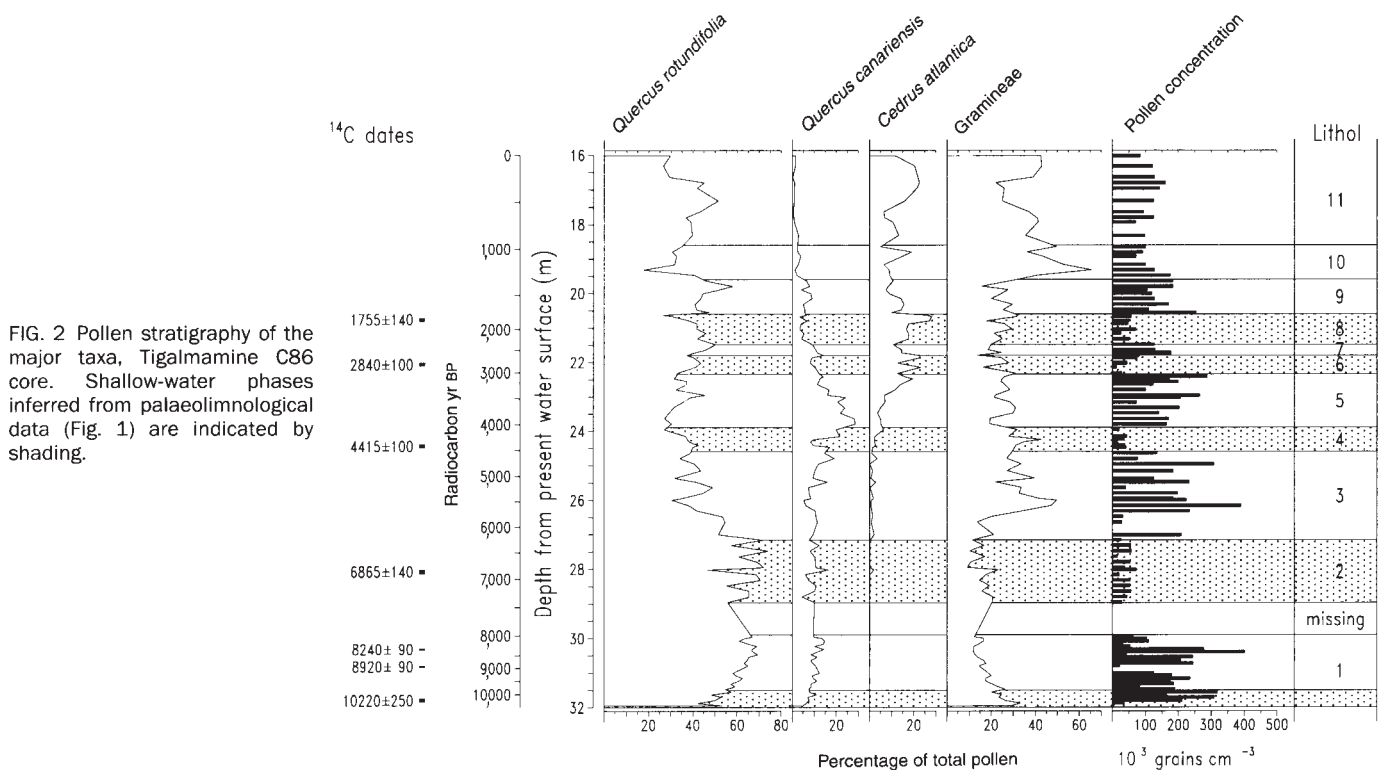


FIG. 2 Pollen stratigraphy of the major taxa, Tigalmamine C86 core. Shallow-water phases inferred from palaeolimnological data (Fig. 1) are indicated by shading.

occurrence and consequences of future climate changes than the gradual, long-term variations³⁴ attributed to orbital (Milankovitch) forcing. □

Received 29 April; accepted 28 November 1994.

- Servant, M. & Servant-Vildary, S. in *The Sahara and the Nile* (eds Williams, M. A. J. & Faure, H.) 133–162 (Balkema, Rotterdam, 1980).
- Owen, R. B., Barthelme, J. W., Renaut, R. W. & Vincens, A. *Nature* **298**, 523–529 (1982).
- Gillespie, R., Street-Perrott, F. A. & Switsur, R. *Nature* **306**, 680–683 (1983).
- Talbot, M. R. et al. *Palaeogeogr. Afr.* **16**, 173–192 (1984).
- Richardson, J. L. & Dussinger, R. A. *Hydrobiologia* **143**, 167–174 (1986).
- Haberyan, K. A. *Freshwater Biol.* **17**, 429–436 (1987).
- Street-Perrott, F. A. & Perrott, R. A. *Nature* **343**, 607–612 (1990).
- Gasse, F., Tehet, R., Durand, A., Gilbert, E. & Fontes, J.-Ch. *Nature* **346**, 141–146 (1990).
- Gasse, F. & Fontes, J.-Ch. in *The Last Deglaciation: Absolute and Radiocarbon Chronologies* (eds Bard, E. & Broecker, W. S.) 295–325 (Springer, Berlin, 1992).
- Gasse, F. & van Campo, E. *Earth planet. Sci. Lett.* **126**, 435–456 (1994).
- Martin, J. *Le Moyen Atlas Centrale: Étude Géomorphologique* (Notes et Mém. Vol. 258, Editions du Service Géologique, Rabat, 1981).
- Lamb, H. F., Eichler, U. & Switsur, V. R., *J. Biogeogr.* **16**, 65–74 (1989).
- El Hamouti, N., Lamb, H. F., Fontes, J.-Ch. & Gasse, F. *C. R. hebdom. Séanc. Acad. Sci., Paris Sér. II* **313**, 259–265 (1991).
- Fontes, J.-Ch. & Gasse, F. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **84**, 191–215 (1991).
- Benkaddour, A. thesis, Univ. Paris XI (1993).
- Fontes, J.-Ch., Coque, R., Dever, L., Filly, A. & Mamou, A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **43**, 41–62 (1983).
- Henderson, P. A. *Freshwater Ostracods* (Universal Book Services/Backhuys, Oegstgeest, 1990).

- Bronshstein, Z. S. *Fresh-water Ostracoda* (Acad. Sci. USSR, Moscow, 1947); also (Amerind, New Delhi, 1988).
- Kröelin, S. & Soulié-Marsche, I. *Quat. Res.* **36**, 210–223 (1991).
- Soulié-Marsche, I. *Bull. Soc. géol. Fr.* **164**, 123–130 (1993).
- El Hamouti, N. thesis, Univ. Paris XI (1989).
- Chivas, A. R., De Deckker, P. & Shelley, J. M. G. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **54**, 43–61 (1986).
- Chivas, A. R., De Deckker, P. & Shelley, J. M. G. *Hydrobiologia* **143**, 135–142 (1986).
- Engstrom, D. R. & Nelson, S. R. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **83**, 295–312 (1990).
- COHMAP members *Science* **241**, 1043–1052 (1988).
- Deil, U. *Rev. géogr. Maroc* **12**, 17–32 (1988).
- Broecker, W. S., Bond, G., Klas, H., Bonani, G. & Wolfli, W. *Paleoceanography* **5**, 469–477 (1990).
- Duplessy, J. et al. *Nature* **358**, 485–488 (1992).
- Bard, E. et al. *Nature* **328**, 791–794 (1987).
- Gasse, F. et al. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **60**, 1–46 (1987).
- Bryson, R. A. *Clim. Dyn.* **3**, 169–177 (1989).
- Lamb, P. J. & Pepler, R. A. in *Teleconnections Linking Worldwide Climatic Anomalies* (eds Glantz, M. H. et al.) 121–189 (Cambridge Univ. Press, 1991).
- Roberts, N., Lamb, H. F., El Hamouti, N. & Barker, P. in *Environmental change in Drylands* (eds Millington, A. C. & Pye, K.) 163–175 (Wiley, Chichester, 1994).
- Kutzbach, J. E. & Street-Perrott, F. A. *Nature* **317**, 130–134 (1985).
- Stuiver, M., & Reimer, P. J. *Radiocarbon* **35**, 215–230 (1993); CALIB rev. 3.0.3.b (Dept Geological Sciences, Univ. Washington, Seattle, 1994).

ACKNOWLEDGEMENTS. This Letter is dedicated to the memory of Jean-Charles Fontes, who played a major role in its genesis. We thank A. Merzouk for logistical support, I. Laidlaw, C. Clarke, R. Maxted, G. Higgs and P. Barker for field assistance, and D. Engstrom for discussions. This work was supported by the Leverhulme Trust and by the UK Natural Environment Research Council.

Effect of bed morphology on flow mixing length at river confluences

Jocelyn M. Gaudet & André G. Roy

Département de Géographie, Université de Montréal, CP 6128, succ. Centre-Ville, Montréal, Québec H3C 3J7, Canada

MIXING processes at river confluences have an important bearing on problems such as pollutant dispersal and the management of river systems, but remain poorly understood. Previous studies¹⁻¹¹ have indicated that flow mixing downstream of a confluence is a slow process, typically completed at distances greater than 100 times the channel width. Bed morphology is now emerging as a critical factor: flume experiments have shown that the width to depth ratios¹² and height discordance^{13,14} between the confluent channels should influence mixing rates¹²⁻¹⁴. But the magnitude of these effects in real rivers is not known. Here we report measurements from three river confluences of moderate size which show that bed discordance can markedly increase mixing rates. For the rivers studied, mixing is complete at distances five to ten times shorter than those reported previously.

This work was prompted by the need to verify experimentally the hypothesis that bed discordance enhances flow mixing rates significantly. To show this effect, three river confluences with channel widths of 5–15 m were surveyed 19, 2 and 2 times, respectively, for conditions varying from low water to the bank-full stage. The waters from the upstream channels of these three confluences differed markedly in their total ion content. These three sites are characterized by a bed discordance between the tributary and main channel, and by the absence of a marked scour hole¹⁵; the height difference between the beds relative to mean flow depth is more important at low water. At each site, water was sampled at numerous and precise points in each channel upstream from the confluence and at six downstream cross-sections located at distances (in units of channel width) of 0.5–60 downstream from the apex of the confluence. Because channel width increases with discharge, distance expressed in channel-width units varies from one survey to the next. Electrical conductivity measures the total concentration of charged ionic species in solution, and it was chosen as an indicator of mixing because of its conservative behaviour¹⁶. (At each site, electrical conductivities are found to be higher in the deeper main stream channel.) The extent of mixing can then be determined from measuring positive and negative percentage deviations of the electrical conductivity from the value expected of the completely mixed flows.

Mixing between the confluent rivers studied here is almost always completed before a downstream distance of 25 channel widths (Fig. 1). This distance is much shorter than previously published mixing lengths of more than 100 channel widths^{1,2,4,5,7,9,11}. As dimensionless distance increases, deviation from complete mixing decreases rapidly. Over a distance of 25 channel widths, deviation values are within the minimal detection-limit interval, indicating that flow mixing is complete. For higher discharges, deviations from complete mixing are initially larger than at lower discharges, and in one occasion mixing was not completed at a distance of 40 channel widths. Because for a given confluence the width/depth ratio decreases with increasing discharge, these results disagree with published data showing that mixing is slower when the width/depth ratio is large¹². In our study, complete mixing is achieved more rapidly at high width/depth ratios (≥ 20). The increased depth differential associated with low water may explain this relationship. A change in relative discharge between the confluent channels has no effect on mixing completion at our sites.

Figure 2 illustrates the effect of depth differential on flow mixing completion, for the site sampled 19 times under different discharge conditions. As bed discordance becomes more important, deviation from complete mixing tends to be smaller. This effect is present in both positive and negative deviation values, but it is more important on the side of the deepest channel, associated with positive deviation values. The effect of bed discordance decreases rapidly with distance downstream. For cross-

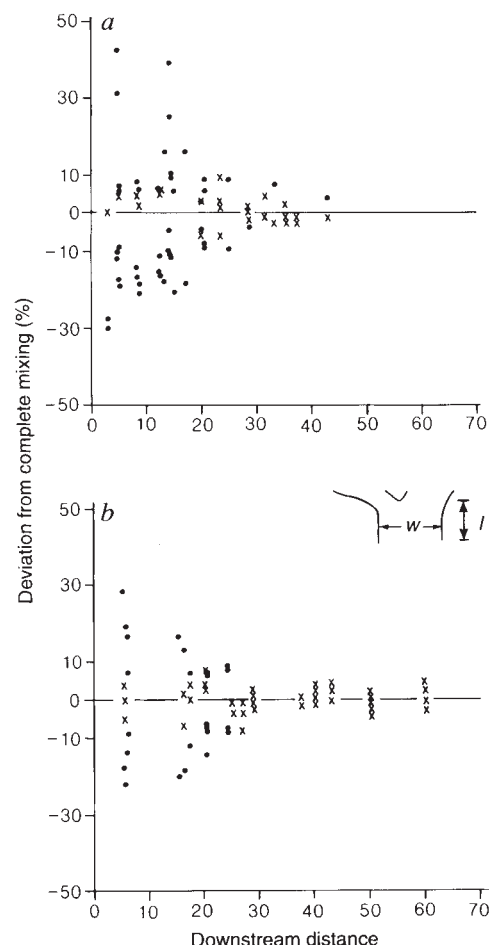


FIG. 1 The relation between mixing completion and dimensionless distance downstream from the apex of the confluence. Vertical axis, the percentage deviation between the observed electrical conductivity and the expected value if mixing was complete; $\frac{([\text{observed}] - [\text{expected}])}{[\text{expected}]} \times 100$. Horizontal axis, the dimensionless downstream distance, defined as downstream sampling distance from the apex (l , in m) divided by the channel width measured 15 m from the apex (w , in m) (see inset). a, Cases where the width/depth ratio is < 20 ; b, cases where this ratio is ≥ 20 . This limit of 20 is the value at which the smaller tributary is half the depth of the main tributary. The width/depth ratio is linked with discharge: values < 20 correspond to high discharges, and values ≥ 20 to low discharges. The width/depth ratio is calculated as the channel width, as defined above, divided by the mean depth of the stream at the same location. The highest positive and negative deviation values observed at cross-sections 3 to 6 (see text) for each sampled event are plotted on this graph, making it a conservative representation. Crosses and dots are deviation values within and without (respectively) the minimal detection-limit interval¹⁹. The values in the minimal detection-limit interval are not statistically distinguishable from perfectly mixed waters at a 95% level of confidence. This figure shows that (1) deviation from complete mixing is negligible from a distance of 25 channel widths, as the data points are in the minimal detection-limit interval; (2) deviations from complete mixing are less important at high width/depth ratio, thus indicating that mixing is more intense.

sections 3 (and 4), located 4.7–5.7 (and 12.6–24) channel widths downstream from the apex, deviation from complete mixing increases as relative bed discordance decreases; deviation values are higher for larger values of depth ratio. For cross-section 5, located 23.5–40 channel widths downstream (not illustrated), this effect is less perceptible because mixing is often completed. At cross-section 6, located 31.6–60 channel widths downstream, mixing is always completed.

The explanation of the effect of bed discordance (and thus of our short observed mixing lengths) is represented in Fig. 3. When water levels are high and bed discordance less important, flows from each confluent channel tend to be segregated along their respective side of the receiving channel. Mixing length in this case is longer, and the observed deviations from complete mixing will be high. When water levels are low and bed discordance relatively more important, water from the shallower tributary tends to flow above the water from the deeper main channel, thus initiating intense mixing very early within the confluence. This overriding effect causes a distortion of the mixing layer and a lateral spreading of the water from the shallower tributary, as previously observed in a laboratory experiment¹³, and is linked with very high turbulence intensities¹⁷. As a consequence, mixing is completed at a similar rate to mixing along vertical profiles. Complete lateral mixing takes place at the shortest distance required to obtain a vertically uniformly mixed flow⁵, explaining the very short mixing distances observed in our experiments.

Similar processes were observed at a large-scale confluence along the St Laurent river¹⁸, showing the range of spatial scales over which this process can take place, from laboratory flume to very large natural confluences. These results are important to river managers and engineers, as many physical and chemical

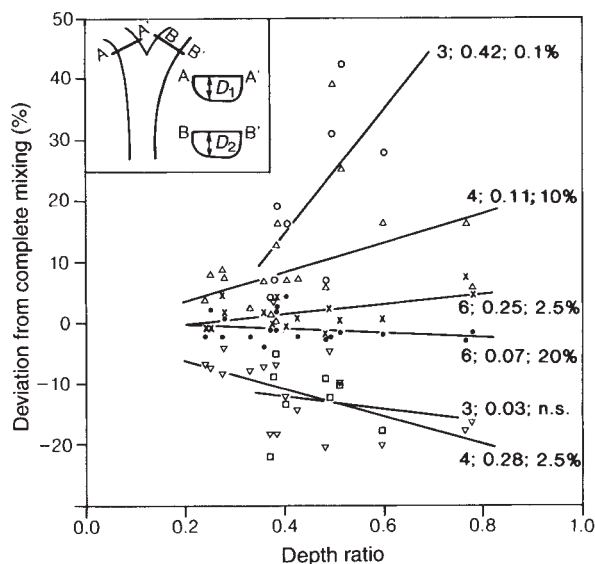


FIG. 2 The effect of the depth differential between the confluent streams on the mixing completion rates, for the site sampled 19 times. Vertical axis, deviation from complete mixing; horizontal axis, depth ratio between the two tributaries for each sampled event, defined as the mean depth of the tributary divided by the mean depth of the main stream ($D_r = D_1/D_2$) (see inset for definition of depth ratio). The highest positive and negative deviation values observed at cross-sections 3, 4 and 6 for each sampled event are plotted, and linear regression lines fitted to the data, for each cross-section and sign of deviation. Values on the graph are the cross-section number, r^2 and the probability (%) that the slope of the relationship is not significant. As the depth ratio increases, deviation from complete mixing is generally greater. This relationship is valid for cross-sections 3 and 4. At cross-sections 5 (not illustrated, for clarity) and 6, deviation from complete mixing is negligible.

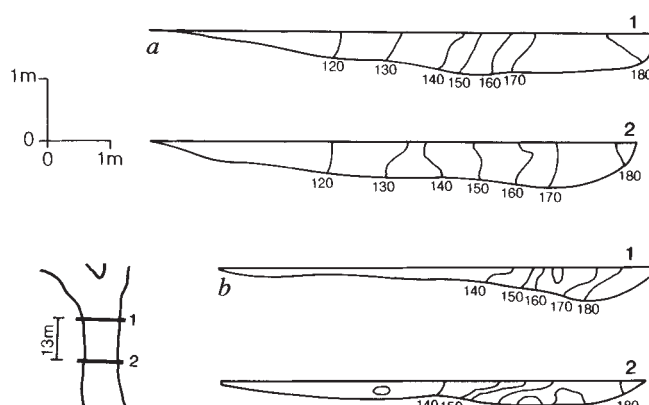


FIG. 3 Summary of flow mixing characteristics immediately downstream from the confluence, at cross-sections 1 and 2 (see diagram at lower left) of the confluence sampled 19 times. Isolines are inferred from mean electrical conductivity values in $\mu\text{S cm}^{-1}$. At high discharges (a), confluent streams are of almost similar depths and relative bed discordance is low: isolines are vertical, showing that the bodies of water from each stream are flowing beside one another. At low discharges (b), water from the shallower channel is clearly flowing above water from the deeper channel as a result of the increased depth differential. This causes a lateral spreading of the shallower-channel waters across the receiving stream, and will eventually decrease mixing length.

variables (such as suspended sediment load, pH, water temperature and chemical concentration and composition) are modified by the nature and intensity of the mixing taking place at confluences. For instance, our findings may be used (1) to identify confluences in a drainage basin with varying mixing completion lengths, (2) to help in positioning effluents in rivers to ensure a fast dispersion of contaminants, and (3) to improve the design or maintenance of channel systems, as the depth differential between tributaries may be used to increase or to decrease the intensity of mixing processes. The results presented here may also help to improve effluent-dispersal models, as most of them do not take into account the effect of the depth differential between confluent streams. □

Received 18 July; accepted 28 November 1994.

- Mackay, R. J. *Can. J. Earth Sci.* **7**, 111–124 (1970).
- Yotsukara, N. & Cobb, E. *Prof. Pap.* 582-C (US Geol. Surv., Washington DC, 1972).
- Engmann, J. E. O. & Kellerhals, R. *Wat. Resour. Res.* **10**, 775–784 (1974).
- Yotsukara, N. & Sayre, W. *Wat. Resour. Res.* **12**, 695–704 (1976).
- Fischer, H. B., List, E. J., Kom, R. C. Y., Imberger, J. & Brooks, N. H. *Mixing in Inland and Coastal Waters* (Academic, New York, 1979).
- Elhadi, N., Harrington, A., Hill, I., Lau, Y. L. & Krishnappan, A. B. *Can. J. Civ. Engng* **11**, 585–609 (1984).
- Bencala, K. E., McKnight, D. M., Zellweger, G. W. & Goad, J. *Water-Sup. Pap.* 2310 (US Geol. Surv., Washington DC, 1987).
- Bencala, K. E., McKnight, D. M. & Zellweger, G. W. *Wat. Resour. Res.* **23**, 827–836 (1987).
- Stallard, R. F. *Wat. Resour. Res.* **23**, 1977–1986 (1987).
- Smith, R. J. *Fluid Mech.* **187**, 589–597 (1988).
- Davis, P. M. thesis, Univ. East Anglia (1991).
- Chu, V. T. & Babarutsi, S. J. *hydraul. Engng* **114**, 1257–1274 (1988).
- Best, J. L. & Roy, A. G. *Nature* **350**, 411–413 (1991).
- Wood, I. R. & Liang, T. J. *Hydraul. Res.* **27**, 587–601 (1988).
- Biron, P., Roy, A. G. & Boyer, C. *Geomorphology* **8**, 115–129 (1993).
- Gaspar, E. in *Modern Trends in Tracer Hydrology* (ed. Gaspar, E.) 1–49 (CRC, Boca Raton, 1987).
- Biron, P., De Seues, B., Roy, A. G. & Best, J. L. in *Perspectives on Flow and Sediment Transport* (eds Clifford, N. J., French, J. R. & Hardisty, J.) 197–213 (Wiley, Chichester, 1993).
- Roy, A. G., Courchesne, F., Giguère, N., Gaudet, J. & Chaumont, D. in *Proc. Joint ASCE & CSCE Nat. Conf. on Env. Engng* (eds Yong, R. N., Hadjicicolaou, J. & Mohamed, A. M. O.) 637–645 (Geotechnical Center, McGill Univ., Montréal, 1993).
- Estimation of Analytical Method Detection Limits* (Municipal/Industrial Strategy for Abatement, Ministry of Environment, Toronto, 1988).

ACKNOWLEDGEMENTS. We thank the many people who helped us in the field or in the laboratory. This work was supported by Fonds FCAR and NSERC.

Trade-off in production between adjacent seamount chains near the East Pacific Rise

Yang Shen^{*†}, Daniel S. Scheirer^{‡§},
Donald W. Forsyth^{*} & Ken C. Macdonald[‡]

^{*} Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

[‡] Department of Geological Sciences, University of California, Santa Barbara, California 93106, USA

BATHYMETRIC data and side-scan sonar images collected near the southern East Pacific Rise between 15° and 19° S^{1,2} have revealed many seamounts on the west flank of the rise, most of which are organized into chains running perpendicular to the rise axis. The number of seamounts of >5 km diameter is at least twice that expected for average East Pacific sea floor³. The formation of numerous, small (<2,500 m in relief) seamounts in ocean basins has been a subject of study for many years^{3–10}, yet we have little understanding of how magma is generated and supplied to build the seamounts. Here we report that the production of seamounts in closely spaced (20–25 km), adjacent seamount chains is negatively correlated, suggesting that the chains share common sources. The existence of closely spaced, long-lived (>5 Myr), linear seamount chains and the tendency for fresh lava flows to be found primarily at the near-ridge ends of the chains suggest that there are plume-like sources in the upper mantle, which are relatively stationary, active for extended periods and affected by adjacent sources.

Most seamounts on the west flank of the southern East Pacific Rise originate within 30 km of the ridge axis¹. In the initial, small study area that we surveyed during December 1990 and January 1991, fresh lava flows are only seen at the axial ends of three major seamount chains (Fig. 1a), indicating that the seamounts are formed near the ridge axis, then passively carried away from the East Pacific Rise by the motion of the Pacific plate at a rate of 70 km Myr⁻¹ (ref. 11). Because the primary seamount building activity occurs from 10 to 50 km off-axis, seamount volume as a function of distance from the ridge may be interpreted to a first approximation as variations in volcanic production or extrusion rate as a function of time. In the initial study area, the largest seamounts are at the western end of the Chapple seamount chain, the centre of the Anakena chain, and the eastern end of the Toroko chain (Fig. 1a). Based on this limited data set, we hypothesized that an increase in the volume of one chain is at the cost of the volume of adjacent chains at similar distances from the axis¹. We suggested that, if this hypothesis holds up with more data, then these seamount chains may share a common source.

Additional bathymetric and side-scan sonar data to test this hypothesis were obtained during November and December of 1992 in an area about 10 times as large as the original survey². We restrict our analysis to the part of the survey area that has nearly continuous bathymetric coverage along the chains, extending from the southern EPR ~340 km off-axis (Fig. 1b). Side-scan sonar images, which have a 20-km wide swath, indicate that no major seamount edifices are missed in gaps of the bathymetric coverage. We divide the seamount field into six equally spaced corridors to minimize subjective grouping of seamount chains. The geometry of the corridors is based primarily on the nearly continuous Cloud chain. In this simple way of dividing

the seamount field, no major seamount chains cross the boundaries of the corridors. In general, seamount volcanism within corridors is more continuous than appears in Fig. 1, with small seamounts or lava flows between major edifices that can be recognized on detailed bathymetry or side-scan sonar maps.

Qualitatively, the additional data appear to support our hypothesis (Fig. 1b): at about 114.5° W, where the Anakena corridor is voluminous, the Toroko and the Chapple corridors have only minor volcanic construction; at about 115° W, where volume in the Toroko corridor is at its peak, there are only minor seamounts in the Chapple and Anakena corridors; further west, where volume in the Chapple corridor is maximum, the Toroko corridor is at a minimum. One exception is at about 116° W where there are voluminous chains in both the Ruru and Chapple corridors. Quantitative analyses (see below), may be used to test whether the apparent negative correlation of volcanism in adjacent chains is statistically significant.

The volumes of seamount edifices within individual corridors fluctuate substantially with distance from the axis (Fig. 2a). Their sum shows a rapid increase in volume in the first 60 km from the ridge axis, then appears to behave as a random deviation from a mean that may increase slightly with crustal age. Side-scan sonar data show that there is minor volcanic activity off-axis (>60 km) in this area that continues to increase the volume of existing edifices, perhaps linking adjacent seamounts in a chain together into ridges². The primary seamount building activity in these chains, however, occurs within 60 km of the axis. To determine whether the observed correlation of volumes of adjacent seamount corridors is significant, we constructed simulated seamount corridors that have the same statistical volume variations as the observed corridors, but which vary independently between each other in volcanic activity. By comparing the observed correlations to correlations of the simulated seamount corridors, we can test the probability that the observed correlations could have arisen by chance. We use simulated corridors for comparison, because the variations in volume for individual corridors are not normally distributed and standard, simple statistical tests based on the assumption of normality cannot be applied directly.

Figure 2b shows an example of a simulated seamount corridor. The simulated corridors share the following statistical characteristics with the real corridors: the ratio of number of positive points to number of zero points, standard deviation, skewness and kurtosis. The only difference between the observed and the simulated corridors is that extrusion rate of each simulated corridor is independent of the others, whereas the extrusion rate of real seamounts may be coherent between adjacent corridors.

We cross-correlated the simulated seamount corridors with each other to obtain a cumulative distribution function (c.d.f.) of the cross-correlation coefficients (Fig. 2c). We then cross-correlated immediately adjacent real seamount corridors and projected the correlation coefficients on the c.d.f. curve (solid dots, Fig. 2c). Four out of five correlation coefficients are negative and fall in the lower half of the distribution; three are in the lowest 3%. The lone instance in the upper half of the distribution is the correlation between the Chapple and the Ruru corridors, the exception pointed out above. The mean of the five coefficients is -0.10, compared to the mean of 0.08 from the randomly varying simulations. A Student's *t*-test indicates that this mean for adjacent corridors is different from the mean for independent corridors at the 99.6% confidence level, and this estimate of confidence level is confirmed by constructing another c.d.f. for the means of groups of five independent correlations.

The cross-correlation of one corridor with another one at least two corridors apart yields three points in the lower half of the distribution and three points in the upper half (open circles, Fig. 2c). Their mean is not significantly different from the mean of the correlations between simulated corridors, suggesting that negative correlation may only exist between adjacent seamount corridors.

[†] Present address: Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.

[§] Present address: Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA.

The statistical tests indicate that negative correlation between volumes of adjacent seamount corridors is significant at a high confidence level. The most probable explanation for such a negative correlation is that adjacent seamount chains share a common source that is roughly constant in volume per unit time¹ (Fig. 2a). Several mechanisms may be responsible for seamount volcanism. They include small-scale, ridge-perpendicular convection^{12,13}, north-south extension of the lithosphere^{14,15}, upwelling associated with plate spreading⁸, preferential melting of heterogeneities embedded within the upwelling mantle beneath the ridge^{6,16}, magmatic solitary waves generated in the primary melt production region¹⁷, and mini-plumes with sources in the deep mantle^{1,7,9}. Stresses within the lithosphere associated with seamount formation may also affect the likelihood that magma can penetrate through adjacent lithosphere^{14,18,19}. The fact that such closely spaced, long-lived, linear seamount chains have not been observed in other ridge-flank surveys²⁰ suggests that their occurrence may be related to a mechanism that is not common to all ridges. Our conclusions apply only to the dominant, near-axis activity in the Rano Rahi field. Other, farther off-axis, volcanism could be caused by other mechanisms^{3,15,21} and, if superimposed on the edifices created near-axis, would degrade the observed negative correlations.

The existence of linear chains, the tendency for fresh lava flows to be found primarily at the near-ridge ends, and the correlation in volume between adjacent chains suggest that there are sources within the mantle that are nearly stationary relative to the ridge, active for an extended period (>5 Myr), and affected by adjacent sources. It is difficult to see how passively embedded heterogeneities could cause coordinated waxing and waning of adjacent chains¹. However, trade-offs in production might arise from plume-forming thermal or chemical instabilities at a boundary layer in the mantle (Fig. 3). We suggest that mini-plumes and trailing conduits developed from the instability. Disturbances generated within the boundary layer repeatedly travel up the conduits, producing fluctuations in discharge within each plume²². As the conduit wave reaches the depth where pressure-release melting begins, a large volume of magma is produced and extracted to form seamounts. Between conduit waves, the magma supply is limited and either no edifices are formed or only small cones are built. Increase in flux to one conduit may drain material from the surrounding source area, causing a decrease in flux to adjacent conduits. With the large number and relatively close spacing of seamount chains, it is unlikely that the instability is at the core-mantle boundary. The instability could be in the upper-mantle transition zone, in the astheno-

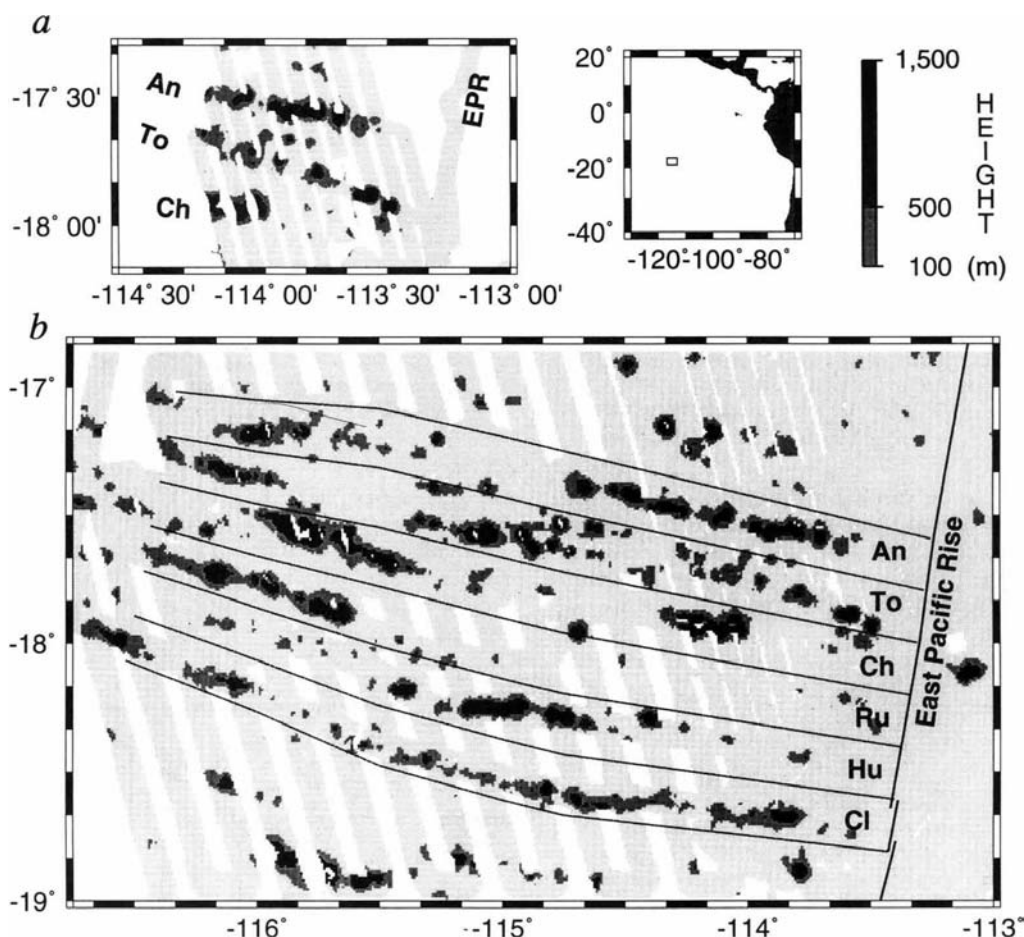


FIG. 1 Distribution of seamounts on the west flank of the East Pacific Rise. Shaded regions indicate continuous bathymetric coverage from SeaBeam2000 multibeam sonar data. Side-scan sonar coverage, from HMR1, is nearly 100%. Darker shades represent the relief of seamounts (m above sea floor, see scale bar) isolated from bathymetric data with an anisotropic median filter¹, which are at least 100 m shallower than the surrounding sea floor. Small-scale map shows location of survey area relative to South America and ridge axes. *a*, Distribution of the three major seamount chains mapped during December 1990 and January 1991. *b*, The Rano Rahi seamount field mapped primarily during

November and December 1992, but including older data as well. The part of the seamount field which has continuous coverage along the seamount chains is divided into six equally spaced corridors, delineated by solid lines. Although some of the seamount chains bifurcate and most of the corridors contain more than one distinct chain, no major groups of seamounts cross the boundaries of the corridors. The Anakena, Toroko, Chapple, Ruru, Hurihuri and Cloud seamount corridors are labelled with two-character symbols standing for the name of the most prominent chain within each corridor.

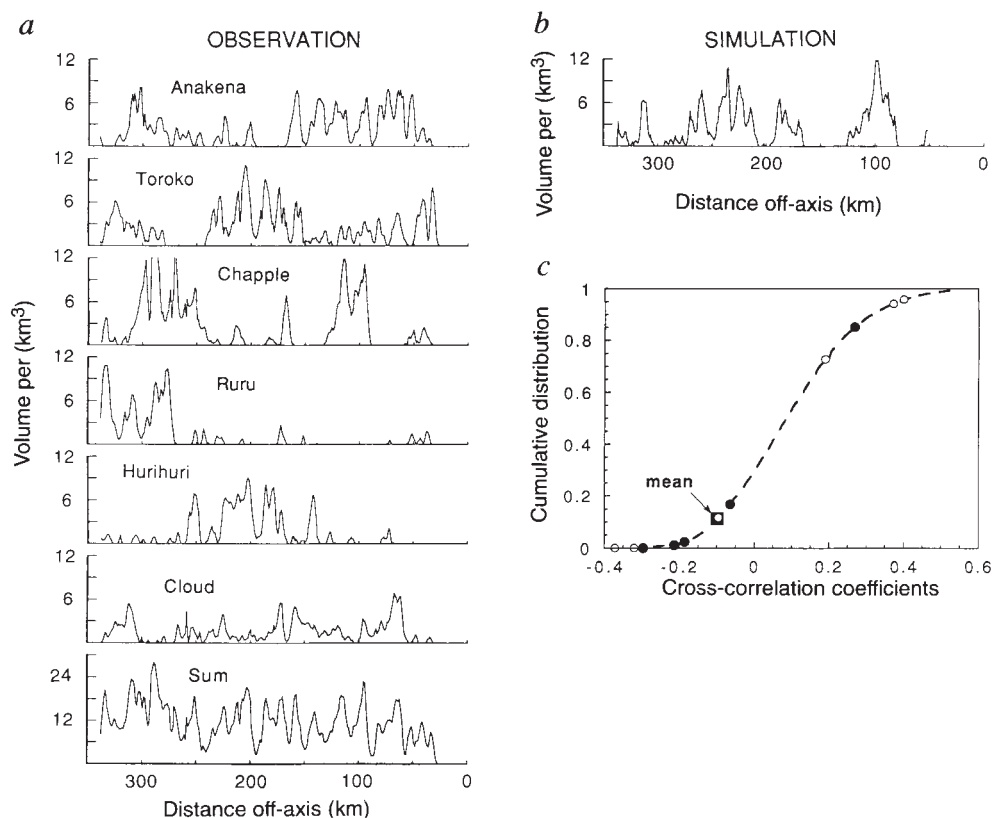


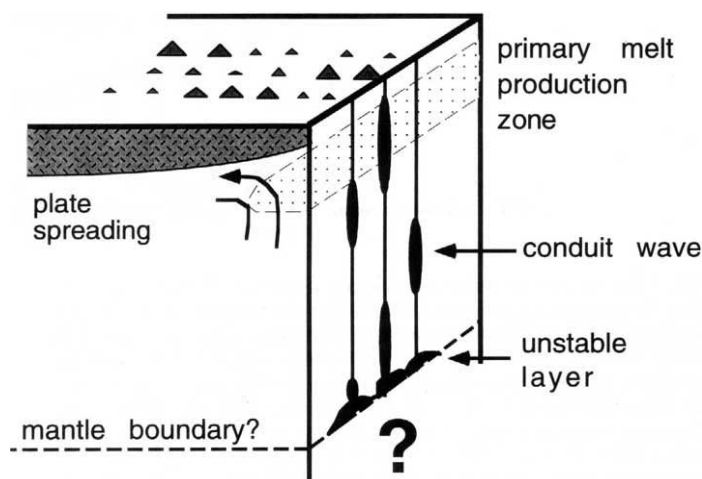
FIG. 2 Comparison of real and simulated seamount chains. *a*, Variation in volume of individual corridors (see Fig. 1) and the sum of the six seamount corridors. Volume per unit distance along the corridors is measured in 1-km bins. *b*, Variation in volume of one of 120 simulated seamount corridors. These were simulated with a fourth-order autoregressive function, $V(t) = a_1V(t-1) + a_2V(t-2) + a_3V(t-3) + a_4V(t-4) + Z(t)$, where $Z(t)$ is a gaussian, random, input variable. V is the volume at time (or, equivalently, distance) t , and the coefficients, a_i , are determined by least-squares fit to the observed, non-zero values of seamount volume at distances >50 km from the ridge axis. Based on partial correlation functions²⁵, higher-order terms are unnecessary. The output from

the autoregressive function is then corrected to the observed cumulative distribution function of volume versus distance, including a non-negative constraint. *c*, Cumulative distribution function (dashed curve) of 7,140 cross-correlation coefficients of simulated seamount corridors. Note that the cross-correlation coefficients are not normally distributed because the variations in volume for individual corridors are not normally distributed. The cross-correlation coefficients of the real, immediately adjacent corridors are projected on the curve as solid dots. Their mean is shown as the solid square. The coefficients of one corridor with another one at least two corridors apart are projected on the curve as open circles.

sphere or at the base of the melting region, although it may be more difficult for sources at shallowest levels to be relatively stationary for an extended period of time in a vigorously upwelling environment. Based on Rayleigh–Taylor instability theory^{23,24} and the spacing between seamount chains (20–

25 km), we estimate that the thickness of the unstable layer is of the order of a few kilometres and the viscosity of the instability layer ($<10^{17}$ Pa s) is much smaller than that of the upper mantle. The unstable layer could also be a potential source for seamounts on the Nazca plate, but with current bathymetric

FIG. 3 Schematic diagram showing preferred model for the formation of the negatively correlated seamount chains. The cross-section parallel to the ridge axis is slightly off-axis. Triangles represent seamounts. Arrow and curves represent upwelling flow associated with plate spreading. The 'primary melt production zone' is the region in which mantle melts to form depleted mid-ocean-ridge basalts. Disturbances generated within a thermal or chemical instability layer travel as waves up the conduit, producing fluctuations in discharges within each plume. Increase in flux to one conduit may drain material from the surrounding region of the unstable layer, causing a decrease in flux to adjacent conduits. Depth of the instability is unknown, although it is not likely to be at the core–mantle boundary.



coverage, it is not clear whether similar chains exist to the east of the East Pacific Rise axis. □

Received 28 July; accepted 22 November 1994.

- Shen, Y., Forsyth, D. W., Scheirer, D. S. & Macdonald, K. C. *J. geophys. Res.* **98**, 17875–17889 (1993).
- Scheirer, D. S., Macdonald, K. C., Forsyth, D. W. & Shen, Y. *Mar. Geophys. Res.* (submitted).
- Smith, D. K. & Jordan, T. H. *J. geophys. Res.* **93**, 2899–2918 (1988).
- Simkin, T. *Mem. geol. Soc. Am.* **132**, 183–193 (1972).
- Fornari, D. J., Ryan, W. B. F. & Fox, P. J. *J. geophys. Res.* **89**, 11609–11083 (1984).
- Davis, E. E. & Karsten, J. L. *Earth planet. Sci. Lett.* **79**, 385–396 (1986).
- Barone, A. M. & Ryan, W. B. F. *J. geophys. Res.* **95**, 10801–10827 (1990).
- Batiza, R., Niu, Y. & Zayac, W. C. *Geology* **18**, 1122–1125 (1990).
- Desonie, D. L. & Duncan, R. A. *J. geophys. Res.* **95**, 12697–12711 (1990).
- Bemis, K. G. & Smith, D. K. *Earth planet. Sci. Lett.* **118**, 251–262 (1993).
- Cormier, M.-H. & Macdonald, K. C. *J. geophys. Res.* **99**, 543–564 (1994).
- Haxby, W. F. & Weissel, J. K. *J. geophys. Res.* **91**, 3507–3520 (1986).
- Buck, W. R. & Parmentier, E. M. *J. geophys. Res.* **91**, 1961–1974 (1986).
- Winterer, E. L. & Sandwell, D. T. *Nature* **329**, 534–537 (1987).
- Sandwell, D. T. et al. *J. geophys. Res.* (submitted).
- Wilson, D. *Earth planet. Sci. Lett.* **113**, 41–55 (1992).
- Spiegelman, M. & Winggins, C. (abstr.) *EOS* **74**, 684 (1993).
- ten Brink, U. *Geology* **19**, 397–400 (1991).
- Macdonald, K. C. *Nature* **331**, 395 (1988).
- Scheirer, D. S. & Macdonald, K. C. *J. geophys. Res.* (in the press).
- McNutt, M. K. & Fisher, K. M. *Seamounts, Islands, and Atolls* 25–34 (Geophys. Monogr. No. 43, Am. Geophys. Union, Washington DC, 1987).
- Olson, P. *Magma Transport and Storage* (ed. Ryan, M. P.) 33–51 (Wiley, New York, 1990).
- Turcotte, D. L. & Schubert, G. *Geodynamics: Applications of Continuum Physics To Geological Problems* (Wiley, New York, 1982).
- Olson, P., Schubert, G. & Anderson, C. *Nature* **327**, 409–413 (1987).
- Box, G. E. P. & Jenkins, G. M. *Time Series Analysis, Forecasting and Control* (Holden-Day, San Francisco, 1970).

ACKNOWLEDGEMENTS. We thank the captain, crew and the scientific parties of the RAPA and Gloria expeditions, and D. K. Smith, R. Batiza and E. E. Davis for reviewing the manuscript. This work was supported by US NSF.

Albanerpetontid amphibians from the Cretaceous of Spain

Gerard McGowan & Susan E. Evans

Evolutionary Anatomy Unit, Department of Anatomy and Developmental Biology (Rockefeller), University College London, Gower Street, London WC1E 6BT, UK

ALBANERPETONTIDS are a group of enigmatic salamander-like fossil amphibians known from deposits of middle Jurassic to Miocene age across Euramerica and Central Asia. Throughout a long history they remained remarkably conservative but can be diagnosed by a suite of unique derived character states, including an anterior peg-and-socket joint between the mandibles, non-pedicellate tricuspid teeth, a distinctive polygonal dermal sculpture pattern, and a two-part craniovertebral joint analogous to that of amniotes. Previous interpretations have placed albanerpetontids within salamanders^{1,2} or as a separate amphibian group^{3,4}. We report here on the recovery of the first complete albanerpetontid specimens (including traces of skin and possible male courtship glands) from the early Cretaceous of Spain. The new material supports the interpretation of albanerpetontids as predominantly terrestrial animals. Albanerpetontids resemble salamanders only in retaining an unspecialized tailed body form; cladistic analysis suggests they represent a distinct lissamphibian lineage.

Albanerpetontids were formally described and named on the basis of disarticulated amphibian bones from the Miocene locality of Grive-St-Alban, France⁵. Although originally linked with the Cretaceous salamander *Prosiren*^{1,6,7}, and often classified within salamanders (Caudata)^{1,2}, some workers have contested this interpretation^{3,4}. Between the earliest record (an atlas from the Middle Jurassic, Bajocian, of France⁸) and apparent extinction in the Miocene, albanerpetontids have been recorded from localities across North America, Europe and Central Asia^{6,9,10}, almost always in the form of dissociated elements. Until now,

the only articulated specimen was a poorly known partial skeleton (the holotype of *Albanerpeton megacephalus*) from the early Cretaceous locality of Pietraroia in Italy⁶.

The early Cretaceous (Barremian) Lithographic Limestone locality of Las Hoyas, near Cuenca, Spain, has yielded a rich freshwater/terrestrial assemblage including plants, aquatic invertebrates, insects, fish, frogs, salamanders, turtles, lizards, crocodiles, dinosaurs and birds^{11–14}. The same deposits have now produced two complete albanerpetontids, one (LH 6020) with exquisite preservation of both hard and soft parts (Fig. 1a–c).

LH 6020 is the skeleton of a small, tailed amphibian of roughly 50 mm snout–vent length. The skeleton is fully ossified and ultraviolet light photography reveals the presence of body scales (Fig. 1a, b). The skin of the head, including the eyelids, contains a mosaic of small polygonal osteoderms, a feature that explains the distinctive skull ornament (Fig. 2a). The structure of the skull itself supports previous reconstructions⁵, but a jugal is recorded for the first time and the slender spike-like form of the cultriform process of the parasphenoid is confirmed. In the axial skeleton, the atlas/axis complex (comprising a bicotylar atlas and a small free axis centrum) is followed by twenty holospondylous presacral vertebrae bearing single-headed ribs, a single sacral vertebra and at least 24 caudals. The scapula and coracoid remain separate and the scapula bears a long anteriorly directed process which braced the limb against the spine. The forelimb is strong, with an ossified radial ball on the humerus and a four-digit manus. In the pelvis, the pubis, ilium and ischium are all ossified. Unfortunately, the hands and feet are contracted (presumably as a result of post-mortem desiccation), making accurate reconstruction difficult. The carpal and tarsal series include three proximal, two or more central and a full set of distal elements. In the hand, the estimated phalangeal formula is 2:3:3:3; in the foot, it approximates 2:3:4/5:3:3.

Within the skin on the anterior surface of each thigh, just proximal to the knee, there is a discrete mass of about 100 small spherules (Fig. 1c). The most reasonable interpretation is of a glandular mass which possibly produced secretions designed to assist grip during courtship. If this is correct, the specimen would be male.

In general, the morphology (tough skin, heavy ossification, powerful limbs and girdles) supports the view⁵ that albanerpetontids were predominantly terrestrial animals.

Comparison of the Las Hoyas albanerpetontid with those of Grive-St-Alban and Pietraroia suggest that the prevailing concept of the genus *Albanerpeton* extending from the Jurassic to the Miocene¹ is incorrect. There is a marked difference between the shape of the frontal in the type species *Albanerpeton inexpectatum*¹ and the related *A. galaktion*³ (strongly triangular, narrow apex, broad, unemarginated orbital margins) and those of the Pietraroia and Las Hoyas specimens (waisted and more rectangular, bulbous apex, curved orbital margins) (Fig. 3a–d). No two species within a modern genus show this degree of difference and it seems a secure basis for generic distinction (other albanerpetontids will be discussed elsewhere). More subtle differences between the Pietraroia and Las Hoyas specimens in the shape of the internasal process and the prefrontal region (see below) suggest that they may be specifically distinct.

Family Albanerpetontidae³
Celtdens gen. nov.

Etymology. *Celtes* (Latin)—chisel; and *dens* (Latin)—tooth; an allusion to the shape of the teeth.

Type species. *Celtdens megacephalus* (= *Albanerpeton megacephalus*⁶).

Generic diagnosis. An albanerpetontid characterized by a waisted frontal with narrow, embayed orbital margins. The internasal process differs from that of *Albanerpeton sensu stricto* in being bulbous rather than spike-like.

Celtdens ibericus new species

Etymology. Ibero—of Spain.

Holotype. LH 6020 (Museo de Cuenca, Cuenca, Spain; provisionally housed in the Unidad de Paleontología, Universidad Autónoma de Madrid, Spain). Part and counterpart of an

almost complete skeleton, missing only part of the palate and tail (Fig. 1a–c).

Horizon and locality. Las Hoyas fossil site, Cuenca Province, Spain. Calizas de la Huérguina Formation (Limestone Unit III). Early Cretaceous (Barremian).

FIG. 1 *Celtdens ibericus*, holotype specimen (LH 6020) photographed under ultraviolet light. a, Complete specimen, $\times 2$; b, head showing osteoderms in eyelids and scale pattern of adjacent skin, $\times 10$; c, left femoral region showing possible glandular mass, $\times 6$.

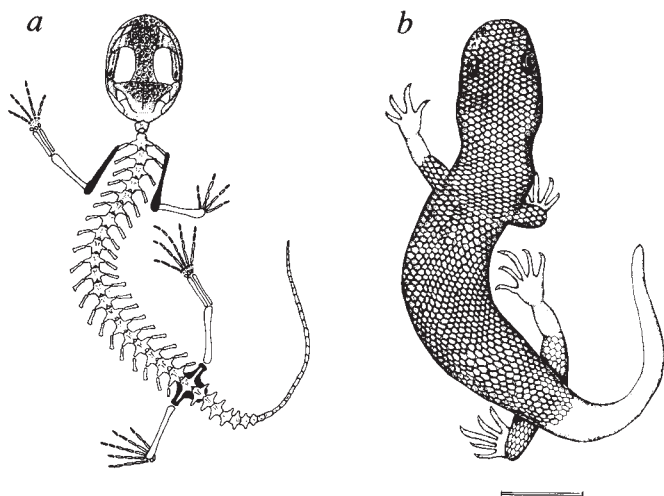


FIG. 2 *Celtdens ibericus*, Las Hoyas, Spain. Reconstructions of a, the skeleton (girdles in black), and b, the body, both in dorsal view. Scale bar, 10 mm.

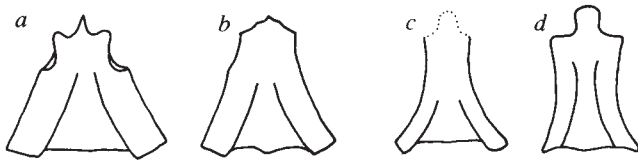
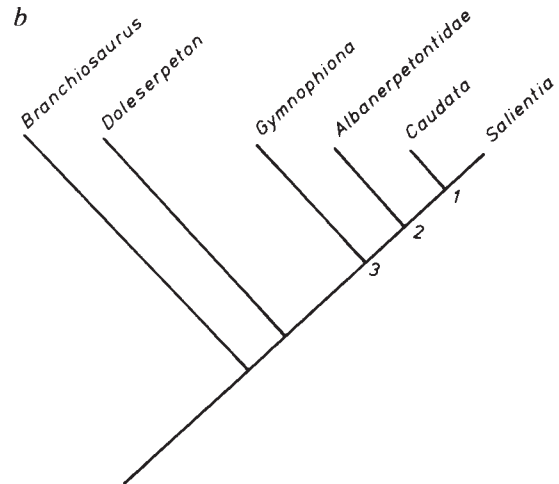


FIG. 3 Frontals of albanerpetontids in ventral view (not to scale): a, *Albanerpeton inexpectatum*, Miocene, France; b, *A. galaktion*, late Cretaceous, Canada; c, *Cetedens megacephalus*, early Cretaceous, Italy; d, *C. ibericus*, early Cretaceous, Spain.

a	1	30
Ancestor	00000 00000 00000 00000 00000 00000	
Branchiosaurus	11100 01010 00000 10000 10000 10011	
Doleserpeton	11110 10021 00000 -0000 00000 01010	
Gymnophiona	11110 10121 1-111 11111 -0000 00001	
Albanerpetontidae	11100 01--2 11111 11100 0-110 -1011	
Caudata	11111 11122 11111 11111 11111 11111	
Salientia	11111 11122 10110 11101 11111 11111	

FIG. 4 Phylogenetic analysis. a, Matrix of taxa versus characters. Character states: 0, primitive state; 1,2, alternative derived states; -, unknown or not applicable. Derived states: 1, short straight ribs; 2, relatively long slender humerus; 3, bicondylar occiput; 4, bicuspid teeth; 5, loss of dermal scales; 6, pedicellate teeth; 7, 22 or fewer presacral vertebrae; 8, loss of palatal fangs; 9, reduction (1) or loss (2) of ectopterygoid; 10, large pleurocentrum, intercentrum reduced (1) or lost (2); 11, neurocranium and dermatocranium form a single structure wrapped around brain; 12, basapophyses present on trunk vertebrae (homologies of gymnophionan parapophyses unclear); 13, radial ball on humerus; 14, loss of tabular; 15, tuberculum interglenoideum present on atlas; 16, loss or reduction of interclavicle; 17, supratemporal lost; 18, postparietal lost; 19, scapula and coracoid coossified; 20, wide parasphenoid; 21, unossified pubes; 22, ossified opercular; 23, postorbital absent; 24, postfrontal absent; 25, jugal absent; 26, subtemporal fenestra separates pterygoid, ectopterygoid (where present) and posterior palatine from jaw margin; 27, frontal enters orbital margin; 28, abbreviated maxilla not contacting jugal or quadratojugal; 29, denticles absent from cultriform process of parasphenoid; 30, jaw articulation anterior to, or level with, occipital condyle. b, Hypothesis of relationships for albanerpetontids. The data matrix was analysed using Hennig 86, resulting in a single tree of 44 steps and a consistency index of 0.72. The outgroups were derived temnospondyl amphibians represented by



Doleserpeton and *Branchiosaurus*^{19,20}. Lissamphibia (node 3) share the derived character states 8,11,13-18; this interpretation requires either independent acquisition of the broad parasphenoid (20) in gymnophionans and Batrachia or secondary loss, with pedicell, in albanerpetontids. The clade including albanerpetontids, caudates and salientians (node 2) is diagnosed by characters 10',12,23,24, requiring a reversal of 12 (basapophyses) in frogs. Batrachia (node 1) share a further six character states: 5,20-22,25,28 not found in albanerpetontids.

Specific diagnosis. As for genus except frontal differs from that of *C. megacephalus* (Fig. 3c,d) in having an interprefrontal width equal, or nearly equal, to the posterior width (in *C. megacephalus*, the interprefrontal width is less than half the posterior width).

The new specimens from Las Hoyas significantly increase our knowledge of this enigmatic amphibian group and permit a more detailed discussion of relationships. Phylogenetic analysis (Fig. 4) indicates that albanerpetontids resemble salamanders only in retaining a primitive tailed body form. The most parsimonious hypothesis places albanerpetontids within Lissamphibia as the sister group of frogs (Salientia) and salamanders (Caudata) (Fig. 4b), although a tree that reverses the positions of gymno-

phionans and albanerpetontids is only slightly longer. The relationships of gymnophionans are currently controversial^{15,16} and resolution of their phylogeny may ultimately affect the lissamphibian cladogram as a whole.

To date, the earliest known lissamphibian is the primitive salientian *Triadobatrachus* from the basal Triassic of Madagascar¹⁷. Although the first records of gymnophionans and salamanders occur later (early Jurassic¹⁵ and middle Jurassic¹⁸ respectively), the three extant lissamphibian groups must have differentiated by the end of the Permian (240 Myr BP). Like them, albanerpetontids had a long independent history. Why the albanerpetontids alone became extinct in the Miocene remains unclear. □

Received 7 September; accepted 11 November 1994.

- Estes, R. *Nature* **224**, 87-88 (1969).
- Trueb, L. & Cloutier, R. in *Origins of the Higher Groups of Tetrapods: Controversy and Consensus* (eds Schultze, H.-P. & Trueb, L.) 223-313 (Cornell University Press, Ithaca and London, 1991).
- Fox, R. & Naylor, B. *Can. J. Earth Sci.* **19**, 118-128 (1982).
- Milner, A. R. in *The Phylogeny and Classification of the Tetrapods* (ed. Benton, M. J.) **1**, 59-102 (1988).
- Estes, R. & Hoffstetter, R. *Bull. Mus. Nat. Hist. Nat. Paris (Sci. Terr.)* **57**, 398-343 (1976).
- Estes, R. in *Handbuch der Paläoherpétologie* (ed. Wellhofer, P.) **2**, (Gustav Fischer, Stuttgart, 1981).
- Duellman, W. E. & Trueb, L. *Biology of Amphibians* (McGraw-Hill, New York, 1986).
- Seiffert, J. *Paläont. Zeitschr.* **43**, 32-36 (1969).
- Nessov, L. *Acta zool. Cracov* **31**, 475-486 (1988).
- Estes, R. & Sanchez, B. *J. Verteb. Paleont.* **2**, 21-39 (1982).
- Sanz, J. L. et al. *Geobios* **21**, 611-635 (1988).
- Sanz, J. L., Bonaparte, J. F. & Lacasa, A. *Nature* **331**, 433-435 (1988).
- Sanz, J. L. & Buscalioni, A. D. *Palaeontology* **35**, 829-845 (1992).
- Perez-Moreno, B. P. et al. *Nature* **370**, 363-367 (1994).
- Jenkins, F. A. & Walsh, D. M. *Nature* **365**, 246-250 (1993).
- Carroll, R. L. & Currie, P. J. *J. Linn. Soc.* **57**, 229-247 (1975).
- Rage, J.-C. & Roček, Z. *Palaeontographica A* **208**, 1-16 (1989).
- Evans, S. E., Milner, A. R. & Mussett, F. *Geobios* **21**, 539-552 (1988).
- Milner, A. R. *Herpetological Monogr.* **7**, 8-27 (1993).
- Boy, J. A. N. *Jb. Geol. Paläont. Abh.* **174**, 75-104 (1987).

ACKNOWLEDGEMENTS. We thank J. L. Sanz and his group at the Universidad Autónoma de Madrid for their support, and A. Diaz-Romeral for collecting and making available LH 6020. Financial assistance was provided by the EC Human Capital and Mobility Program, the Royal Society of London, the Junta de Comunidades de Castilla-La Mancha and an SERC Research Studentship to Gerard McGowan. C. Sym and J. Pendjiky took the photographs.

Structural engineering of an orb-spider's web

Lorraine H. Lin^{*†}, Donald T. Edmonds^{*}
& Fritz Vollrath^{‡§}

^{*} Clarendon Laboratory, University of Oxford, Oxford OX1 3PS, UK

[†] Ove Arup and Partners, 13 Fitzroy Street, London W1P 6BQ, UK

[‡] Department of Zoology, University of Aarhus,
DK 8000 Aarhus C, Denmark

[§] Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

THE silk of spiders first evolved 400 million years ago and orb webs emerged 180 million years ago^{1,2}; present-day spiders' webs are structures efficiently engineered by nature. The planar orb web of the garden spider *Araneus diadematus* has evolved with the prime function of capturing fast-moving and, on a relative scale, massive insects. We have now analysed the structural engineering of a complete web, using computer modelling, to incorporate the measured time-dependent stress-strain characteristics of the two chief types of web silk. With this model we unexpectedly discovered that aerodynamic damping plays a crucial role in prey capture, an observation that we confirmed in laboratory experiments on real webs.

The web of the orb spider *A. diadematus* is mainly composed of two types of thread. There is a framework of dry radial threads upon which is laid a single wet (glue-covered) captured spiral. The stiff radial threads have time-dependent viscoelastic properties, the pliant capture threads have essentially time-independent elasticity³. Although the diameters of the two types of thread are similar their dynamic response to applied stress is strikingly different⁴⁻⁶, which is best shown when measured in a rapid-response microbalance (response time <30 ms) such as that used in a detailed comparative analysis of their stress-strain characteristics³. This balance allows us to measure time dependence of the viscoelastic thread with sufficient accuracy for advanced computer modelling of entire webs. Figure 1 shows a typical stress-strain characteristic of a radial thread. The equations used to represent the two thread types in our model are given in the figure legend.

FIG. 1 The dynamic time-dependent viscoelastic behaviour of a single radial thread taken from a web of *A. diadematus*. The thread was stretched at a strain rate of $dX/dt = 0.01 \text{ s}^{-1}$ for 5 s, allowed to relax for another 5 s and this sequence was repeated twice. Each time, the silk showed a rapid initial increase in resisting force for a rapid change in length, but eventually it relaxed towards its equilibrium tensile force for that extension, as indicated by the broken line. The loading was then repeated with a negative strain rate of $dX/dt = -0.01 \text{ s}^{-1}$; again the silk responded with a low tension, or even sagged, but eventually tightened. The equations used to represent the two types of thread within the computer model were obtained from similar types of measurements on real threads from a single web using our fast-recording microbalance³. The tension T_s exhibited by the sticky spiral thread extended to a length L is a nonlinear but time-independent function of the imposed strain. It can be represented by a simple equation of the form

$$T_s = A(X + BX^n) \quad (2)$$

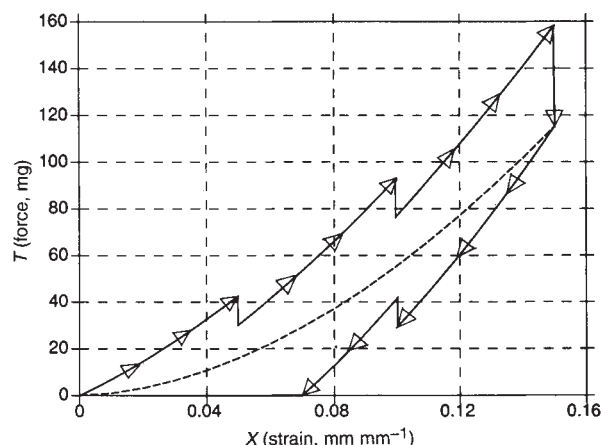
where X is the strain given by $X = (L - L_0)/L_0$; L_0 is the unstretched length and A , B and n are constants chosen to fit the measured stress-strain characteristic of this type of thread (with T_s measured in mg; $A = 27.8 \text{ mg}$, $B = 134$ and $n = 1.85$ for the radial thread; $A = 0.359 \text{ mg}$, $B = 0.00165$ and $n = 9.25$ for the spiral thread). The time-dependent behaviour of the stiff radial thread results in hysteresis in the stress-strain curve and to energy dissipation. It may be simulated by adding a second time-dependent tension $T_2(t)$ to a tension T_1 of the type represented by equation (2), so that the total tension in the radial thread becomes $T_1 + T_2(t)$. The equation used

To build a computer model of the complete web, we have adapted the nonlinear structural analysis (finite element) programming environment OASYS DYNA3D which is used primarily to test 'crash worthiness' in the automobile industry. Our program allows us to simulate in three dimensions the dynamic behaviour of a cable net of arbitrary geometry, in which a time-dependent and nonlinear stress-strain characteristic may be assigned to each cable link between any two adjacent nodes of the net. In this manner a real web may be modelled, incorporating experimentally measured stress-strain characteristics of real threads. Iterating between model predictions and matching experiments on natural webs leads to unforeseen insights into orb web function. These were previously untestable.

Measurements on real webs reveal that the stiff radial threads are pre-tensioned in the resting configuration of the web^{6,7}, and thus dominate the web structure. The much less-stiff spiral threads form a secondary structure laid on the primary structure to constitute an 'anisotropic membrane' with differing radial and circumferential stiffness⁸. Theoretically, and this is confirmed by our model, the forces from the impact of prey follow a predictable path through the web always moving in the direction of greater stiffness. From the point of impact, most probably on the capture spiral⁹, the disturbance moves through the secondary structure to the nearest primary members, then radially outward to the support frame of the web.

The most important new feature of an orb web, as we learned from comparing model predictions and real web experiments, is the importance of aerodynamic damping in the dissipation of an insect's kinetic energy upon impact in a web. The hysteretic behaviour of the radial threads (see for example Fig. 1) as they stretch and relax leads to some energy dissipation^{4,9-12}. However, our simulations and experiments now show that this form of energy dissipation is too weak to account for all of the energy absorbed by a web, and that only by inclusion of additional aerodynamic damping can the behaviour of real webs be reproduced.

What is the mechanism behind our observations? At the microscopic scale of the spider silk, with diameters of about $1 \mu\text{m}$ and velocities of about 1 m s^{-1} , the air flow is laminar, inertial forces may be neglected and viscous forces dominate. In these circumstances^{13,14} the drag force per unit length of thread is proportional to the velocity of motion but is independent of



to represents $T_2(t)$ in the model was:

$$T_2(t) = \exp(-t/\tau) \left[T_2(0) + C \int (dX/dt) \exp(t/\tau) dt \right] \quad (3)$$

where C and τ are constants ($C = 600 \text{ mg}$, and $\tau = 20 \text{ s}$) obtained by comparison with stress-strain experiments on the radial thread and $T_2(0)$ is the tension at the time $t = 0$ when the strain is applied. However, $T_1 + T_2(t)$ must be positive or zero as the thread cannot resist compressive loads.

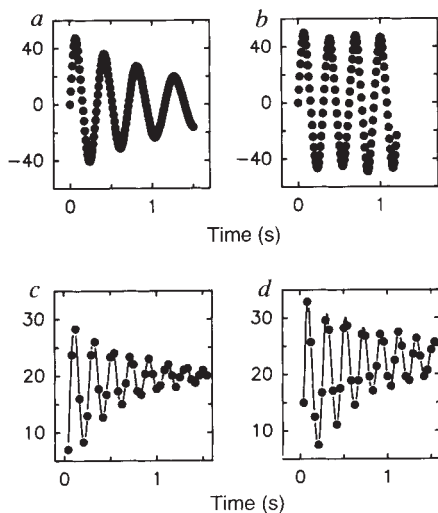


FIG. 2 Comparison of simulation and data of a pellet hitting a spider's orb web. *a*, Output of the computer model showing out of plane displacement of a radial thread at the point of impact of a pellet of 47 mg at a speed of 1.9 m s^{-1} when aerodynamic damping is included in the model (deflections in mm). *b*, The same simulation in which the measured hysteresis of the radials was included, as normal, but aerodynamic damping was omitted. *c*, Laboratory measurement of the out of plane displacement of a captured Styrofoam bullet (at 21 mg, 1 m s^{-1}). *d*, Another measurement on the same web and bullet after all spiral sections had been cut away with a hot wire showing the large reduction in aerodynamic damping. The experimental data of *c* and *d* were recorded on video at 50 frames per s and analysed frame by frame.

its diameter. The architecture of the web takes advantage of aerodynamic damping by providing a fine net that covers a large area but with no penalty for using very fine threads. Model predictions of the out of plane web deflection as a function of time following simulated prey capture are shown in Fig. 2*a, b*. This is with and without the inclusion of aerodynamic damping to illustrate its large effect.

The importance of aerodynamic damping suggested by our simulations was verified by measurements on a vertical, real orb web with which we analysed damping after a deflection perpendicular to its plane. We shot a pellet into the web and recorded the motion of various points on the web on video (at 50 frames per s) for subsequent single-frame analysis. The deflection (D) of an intact web as a function of time t was well represented (Fig. 2*c*) by an exponentially damped sine wave of

the form

$$D = A \exp(-t/\tau) \sin(2\pi ft) \quad (1)$$

where A is a constant. The exponential decay time (τ) was 0.49 s and the frequency of vibration (f) was 5.41 Hz. All spiral thread sections were then cut away between the radials using a hot wire and the displacement behaviour observed (Fig. 2*d*) again. Once more, the displacement was well represented by the function in equation (1) but now with $\tau = 0.98 \text{ s}$ and $f = 5.40 \text{ Hz}$.

The increase in damping by a factor of about 2 with the spiral present showed the importance of the aerodynamic damping acting on the spiral thread. Calculations revealed that the aerodynamic damping of the spiral is enhanced by about 35% by the presence of the glue droplets. At the small strains imposed on the spiral threads in this experiment, they were perfectly elastic

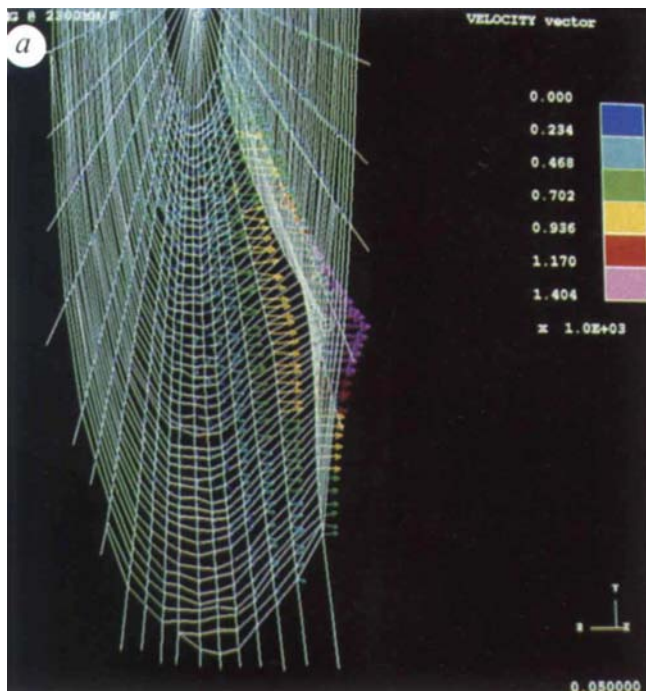
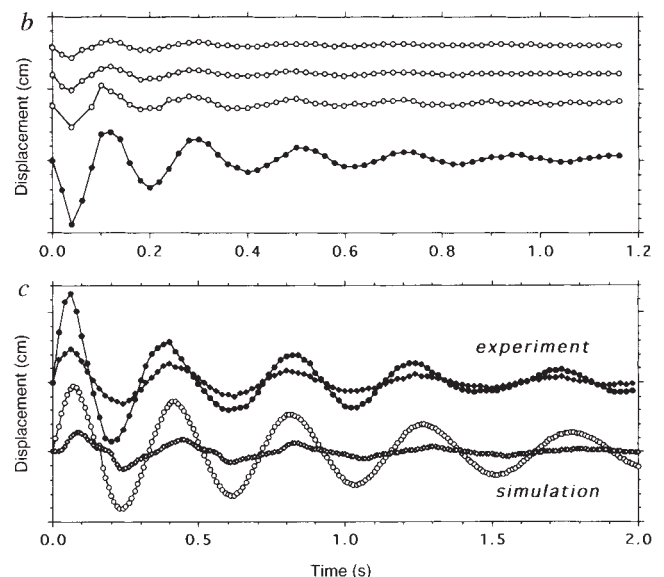


FIG. 3 Stability of a typical *A. diadematus* orb web after prey capture. *a*, The predicted initial velocities of the nodes of a web struck by a pellet of 8.2 mg with an initial speed of 2.3 m s^{-1} . The velocities are colour-coded according to the scale at the right of the diagram. *b*, Measurement (see Fig. 2) on a real web of the deflection at the point



of impact and at the three next neighbouring radial threads at the same distance from the hub. The pellet had the weight and initial velocity used in the simulation shown in *a*. Deflection is confined to close neighbouring radials whereas the bulk of the web remains essentially motionless as predicted by the model in *a*. *c*, A comparison between prediction and measurement of the deflection at the point of impact and at the web hub after the impact of a pellet of 47 mg and initial velocity 1.9 m s^{-1} . The pellet impacted at a distance of 0.11 m from the centre of a web of diameter $\sim 0.3 \text{ m}$.

and hence did not contribute to the hysteretic damping of the web. The fact that the frequency of vibration was unchanged when the spiral thread was removed is consistent with the hypothesis that the major restoring forces are due to the radial threads.

Radial pre-tension and web geometry confine web deflection after prey capture to a sector of the web near the capture site (Fig. 3). Figure 3a shows the computed initial velocities of all the nodes of a portion of a web after capture of a pellet. From the model it is easy to obtain a continuous record of the stresses and strains in all the radial and spiral threads which is not possible with a real web. Figure 3b shows the measured deflection of a real web under the impact of the same pellet measured at the capture site and at the first, second and third neighbouring radial threads at the same distance from the hub. The motion is seen to have diminished to near zero at a site three radial spokes away from the capture site. The reliability of the model is further illustrated in Fig. 3c which shows the close similarity between the predicted and measured deflection of the capture site and the web hub after capture of identical pellets. The aerodynamic behaviour of a web presents a constraint as well as opportunities.

Figure 4a shows the predicted shape of a web subjected to wind at constant velocity perpendicular to its plane. Figures 4b, c show good agreement between the perpendicular deflection along a diameter of a web at wind speeds of 1.01, 2.67, 5.00 and

6.77 m s^{-1} measured on a real web and predicted by the model. Owing to the extreme anisotropic stiffness of the web, the deflection is largely perpendicular to the web for all wind directions. Thus, aerodynamic drag improves prey capture but carries with it an increased risk of web destruction by winds. Web loss is serious for the spider as the web is normally ingested to recycle the silk¹⁵ and also to harvest some of the water absorbed from the air¹⁶ by the hygroscopic glue. This may be one reason orb webs are often oriented with the normal to their plane at an angle to the primary wind direction¹⁷.

These results are a step towards the important goal of understanding the great diversity of spider webs, their structures and building materials. It opens up the field of animal engineering illustrating the balance between costs and benefits and also shows the effects of scale in an exemplary way. □

Received 16 June; accepted 15 November 1994.

1. Shear, W. A., Palmer, J. A., Coddington, J. A. & Bonamo, P. M. *Science* **246**, 479–481 (1989).
2. Selden, P. A. *Nature* **340**, 711–712 (1989).
3. Vollrath, F. & Edmonds, D. T. *Nature* **340**, 305–307 (1989).
4. Denny, M. J. *J. exp. Biol.* **65**, 483–506 (1976).
5. Gosline, J. M. et al. in *Material Science and Biotechnology* (eds Kaplan, D., Adams, W. W., Farmer, B. & Viney, C.) 328–341 (Am. chem. Soc., Washington, 1994).
6. Kohler, T. & Vollrath, F. *J. exp. Zool.* **270**, 1–17 (1994).
7. Wirth, B. & Barth, F. G. *J. comp. Physiol.* **a171**, 359–371 (1992).
8. Lakes, R. *Nature* **361**, 511–515 (1993).
9. Rhisiart, A. & Vollrath, F. *Behav. Ecol.* **5**, 280–287 (1994).
10. Langer, R. *Am. Zool.* **9**, 81–89 (1969).
11. Gosline, J. M., DeMont, M. E. & Denny, M. W. *Endeavour* **10**, 37–43 (1986).
12. Craig, C. L. *Biol. J. Linn. Soc.* **30**, 135–162 (1987).
13. Goldstein, S. *Modern Developments in Fluid Mechanics* (Dover, New York, 1955).
14. Massey, B. S. *Mechanics of Fluids* 5th edn (Van Nostrand Reinhold, London, 1987).
15. Breed, A. L., Levine, V. D., Peakall, D. B. & Witt, P. N. *Behaviour* **23**, 43–60 (1964).
16. Edmonds, D. T. & Vollrath, F. *Proc. R. Soc. B248*, 145–148 (1992).
17. Hieber, C. S. Z. *Tierpsychol.* **65**, 250–260 (1984).

ACKNOWLEDGEMENTS. Much of this research was made possible by J. Zunz, T. Barker and the late P. Rice of Ove Arup Partnership, London. We also thank C. Wood of the Engineering Department, Oxford University, and A. Allsop of Ove Arup Partnership for advice on the aerodynamic drag calculation and R. Plano and S. Ishida of the Natural Structures Laboratory of Renzo Piano Building Workshop, Genoa who supported this work.

Lek size, male mating skew and the evolution of lekking

Fredrik Widemo*† & Ian P. F. Owens†

* Section of Animal Ecology, Department of Zoology, Uppsala University, Villavägen 9, S-752 36 Uppsala, Sweden

† Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY and The Galton Laboratory, Department of Genetics and Biometry, University College London, London NW1 2HE, UK

DESPITE extensive theoretical effort^{1–8}, the evolution of lekking as a mating system remains a controversial issue^{9,10}. Leks are non-resource-based mating aggregations², but may also be regarded as patches differing in female encounter rate^{2,3,5,7}. We report here a new distribution model that incorporates variation in male mating skew with lek size. The model predicts that, under specified conditions, high-ranking males have smaller optimal lek sizes than low-ranking males. All males benefit from initial clustering, but only low-ranking males gain from large aggregations. This generates progressive clustering around high-ranking males at hotspots determined by female spatial distributions. The predictions of our model were validated in two ways using empirical data on lekking ruffs, *Philomachus pugnax*. Our model integrates the basic elements of the previously competing hotspot^{2,3} and hotshot⁴ models of lek evolution by a simple mechanism, and could explain the evolution of lekking.

† To whom correspondence should be addressed.

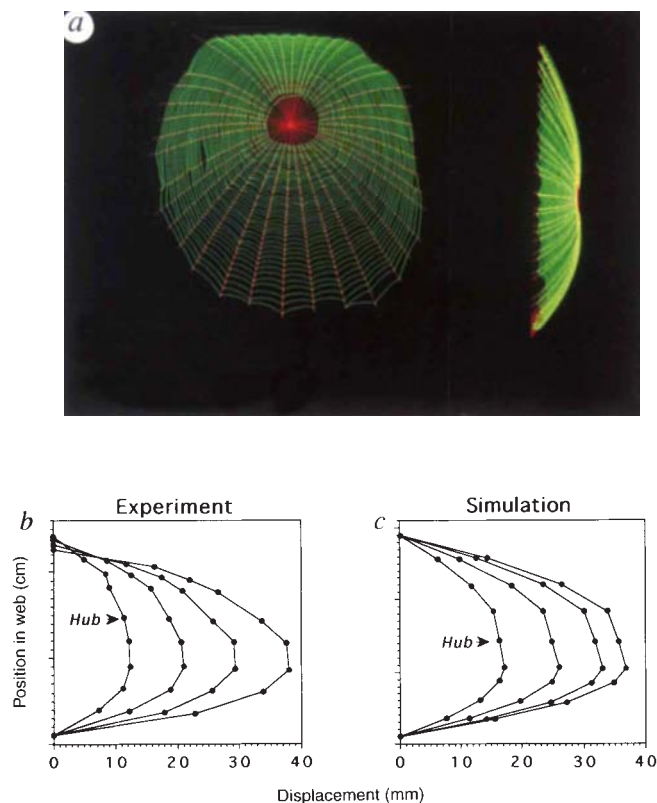
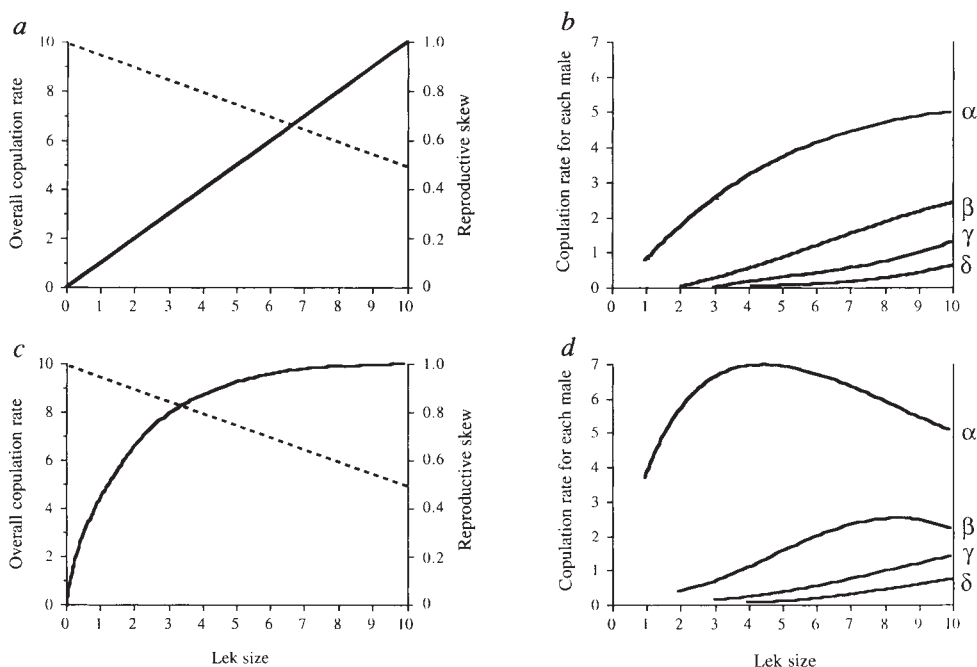


FIG. 4 Wind loading of an *A. diadematus* web. a, Computer simulation of a web billowing in a wind of 10 m s^{-1} directed normal to the plane of the web. b, c, A comparison of the out of plane displacement along two near vertical radials forming a diameter measured (see Fig. 2) on a real web in a wind tunnel and predicted by the model for wind speeds of 1.01, 2.67, 5.00 and 6.77 m s^{-1} . Radial threads break at tensions around $900 \mu\text{N}$, spiral threads at about $40 \mu\text{N}$ (ref. 4). Depending on the size and the orientation of the web, its radials are in danger of breaking near the periphery at wind speeds between 5 and 15 m s^{-1} . The danger is particularly great for winds normal to the web plane because of the extreme anisotropic stiffness of the web.

FIG. 1 Skew model of mating success across males of different rank. *a*, First set of hypothesized relationships between lek size, measured in number of territorial males, and both overall copulation rate and mating skew among males. Overall copulation rate increases linearly with increasing lek size. Mating skew decreases linearly with increasing lek size. *b*, Corresponding predicted copulation rates for four territorial males of decreasing rank. *c*, Second set of hypothesized relationships between lek size and both overall copulation rate and mating skew. Overall copulation rate increases asymptotically with increasing lek size whereas mating skew decreases linearly with increasing lek size. *d*, Corresponding predicted copulation rates for four territorial males. Model: leks consist of a group of males that are visited by females and among whom a simple dominance hierarchy exists. The frequency with which females visit different leks is predicted by the size of the leks. Upon entering a lek a female approaches the highest-ranking male, the alpha male. The alpha male and the female will then copulate with a certain probability. The probability that any pairing results in a copulation is proportional to the mating skew at that lek size. Male mating skew is expressed as the ratio of observed skew to maximum possible skew^{15,16} (only one male copulates: skew=1; all males copulate equally: skew=0). If the alpha male does not copulate



with the female, then the female will approach the next-highest-ranking male, the beta male. The beta male and the female will then copulate with the same probability. This process continues with the female approaching males in descending order of dominance until she either copulates with one of them or until she has approached all the males. The female then leaves the lek.

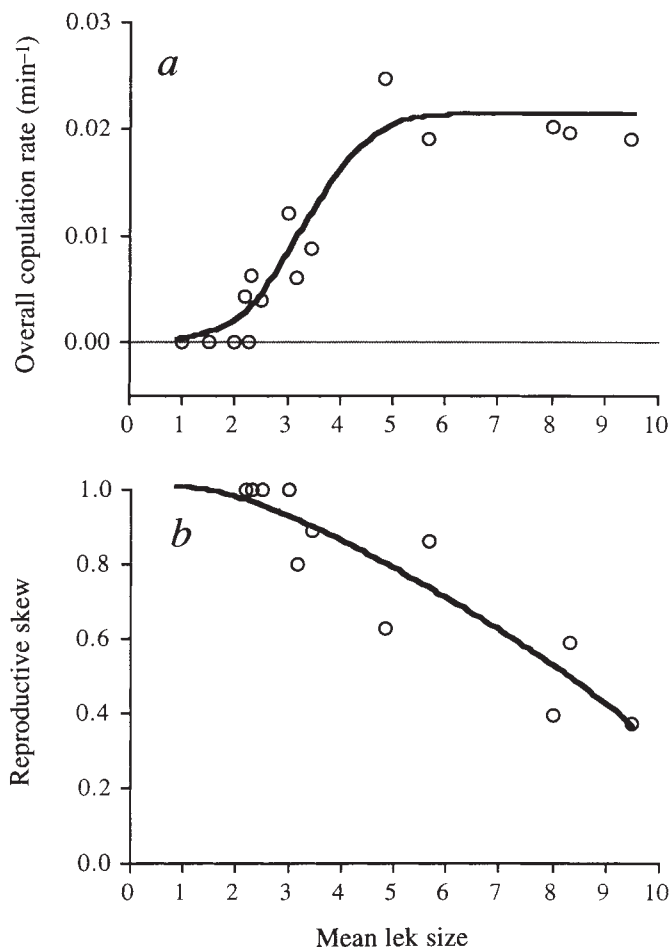


FIG. 2 Observed relationships across leks between mean lek size and *a*, overall copulation rate (logistic regression; 'copulation frequency' = $-0.00008 + 0.02066 / (1 + \exp(-1.842(\text{'mean lek size'} - 3.290)))$, $F_{3,11} = 38.12$, $P < 0.001$) and *b*, skew in male mating success (exponential regression; 'skew' = $1.5419 (0.6629^{\text{'mean lek size'}})$, $F_{1,9} = 43.54$, $P < 0.001$). Mean lek size is the mean number of territorial males on each lek across all observation periods within a season. Overall copulation rate is the total number of copulations by territorial males on a lek in a season divided by the length of time over which the lek was observed. Male mating skew index is calculated as the ratio of observed skew to maximum possible skew^{15,16}: skew index = $(N_b v + N_n) / (N_b + N_n)$ where N_b denotes the number of successful males in the group, N_n the non-successful males, and v represents the variation in copulations among successful males. With a single successful male, $v = 1$; with several successful males $v = N_b (\sum (p_i - 1/N_b)^2) / (N_b - 1)$ where p_i is the proportion of all copulations performed by the i th male. Fifteen leks were observed over four seasons. In each season, data comes from morning observations within 10 days of the annual population mating peak only. To avoid pseudoreplication, only the season in which the mean number of territorial males was the highest was included for each lek. Regression lines were drawn according to a best fit among linear, logarithmic and gompertz models.

We applied distribution theory¹¹ to lekking and developed a model that calculates the expected mating success for males of given rank while varying lek size, overall copulation rate and level of mating skew (Fig. 1). The model is based on empirical studies that have shown male mating skew to be negatively related to lek size^{5,7}, possibly as a consequence of the increased levels of male aggression on larger leks. The interaction between male mating skew and lek size has important implications which hitherto have been overlooked. If increased levels of disturbance affect the abilities of males to monopolize soliciting females, high-ranking males will be more seriously affected by disturbance, in terms of lost copulations, than will low-ranking males. Consequently, conflicts of interest might ensue between males of differing phenotypic quality at certain lek sizes, where the addition of another male may decrease the top males' ability to monopolize matings.

Our model shows that a negative relationship between lek size and male mating skew results in high-ranking males having intermediate optimal lek sizes, unless the effect of the decreasing skew is more than offset by a concomitant increase in overall copulation rate with lek size (Fig. 1a–d). Furthermore, in such cases, low-ranking males are predicted to have larger optimal lek sizes than high-ranking males. This effect is dependent on the relative shapes of the relationships between lek size and overall copulation rate, and mating skew, respectively. It is, therefore, impossible to make inferences about optimal lek sizes for individual males from the relationship between lek size and overall copulation rate alone.

To test the predictions of our model we studied the ruff, a lekking wader, on southern Gotland in the Baltic Sea over four breeding seasons. The overall copulation rate was found to increase sigmoidally with lek size (Fig. 2a). The skew in male mating success, however, decreased exponentially with lek size (Fig. 2b). High-ranking males should, therefore, have intermediate optimal lek sizes, whereas low-ranking males should have higher success on the largest leks. In agreement with the predictions, high-ranking males were found to have smaller optimal lek sizes compared to low-ranking males (Fig. 3a–d).

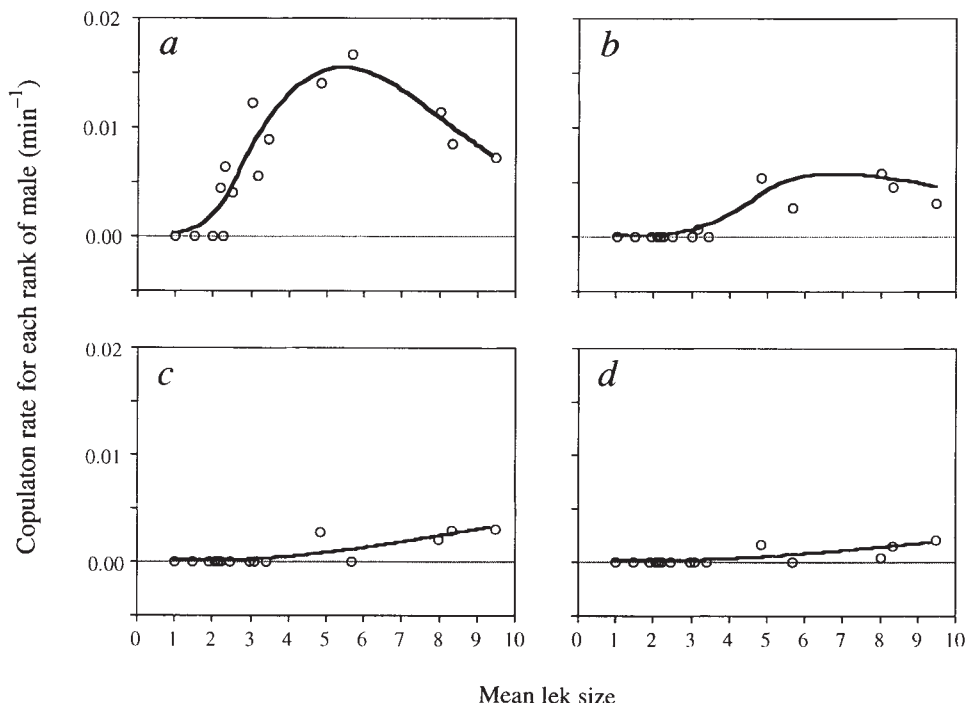
Male ruffs actively entice both females and males to their leks¹² and larger leks attract more females^{7,13}. High-ranking males should, therefore, actively try to increase the lek size up to their mating optimum, whereas low-ranking males should continue actively to attract males beyond this point. As predicted, high-ranking males enticed at a significantly lower rate than lower-ranking males towards flying males at larger lek sizes, whereas there was no difference at smaller lek sizes (Fig. 4). Moreover, the shift in enticement behaviour for high-ranking males coincided with their observed optimal lek size.

Distribution theory in different forms has previously been applied to explain male clustering, and hence, lekking^{2,3,5,7}. It is, however, difficult to see how female distribution patterns alone could lead to the extreme clustering observed in many lekking species. Previous distribution models thus seem insufficient to explain the evolution of lekking.

Both our distribution model and the empirical data show a conflict of interest between males of different rank with respect to lek size. Although earlier models incorporating fixed levels of skew over all lek sizes have shown that increasing skew results in decreased clustering³, our model generates a mechanism that promotes clustering, as low-ranking males gather around high-ranking males at hotspots determined by female distribution patterns. All males initially gain from clustering at those hotspots. As leks increase further in size the majority of the males still benefit from clustering, but the highest-ranking males incur a decrease in fitness. This outcome resembles some of the predictions of the hotshot hypothesis, where unattractive males cluster around attractive males, parasitizing their attractiveness.

High-ranking males do not appear to abandon leks until they get very large, probably because abandonment would be associated with severe costs. Lek sites are highly traditional^{7,14}, and males are more or less restricted to displaying on existing sites. Furthermore, in ruffs, switching lek site is always associated with prolonged periods of increased fighting (our unpublished observation). Top males may, therefore, be effectively trapped in increasingly suboptimal places because of being followed by lower ranking males. Hence, although all males favour the evolu-

FIG. 3 Observed relationships between overall copulation rate for each category of male and mean lek size for a, alpha males, b, beta males, c, gamma males, and d, delta males. Males were ranked by proportion of copulations within a lek or, when the proportion of copulations were equal between males, by proportion of time present on the lek. The relationships between lek size and copulation rates for each rank of male were calculated by multiplying the relationships between lek size and overall copulation rate (Fig. 2b) by the relationship between lek size and the proportion of copulations obtained by each rank of male. The relationships between lek size and the proportion of copulations obtained by each rank of male were determined by regression models; regression lines were drawn according to the best fit among linear, logarithmic and gompertz models (alpha males: exponential regression, $F=15.6$, $r^2=0.78$, $P<0.001$; beta males: logistic regression, $F=13.3$, $r^2=0.71$, $P<0.01$; gamma males: exponential regression, $F=15.38$, $r^2=0.74$, $P<0.001$; and delta males: exponential regression, $F=10.34$, $r^2=0.65$, $P<0.001$).



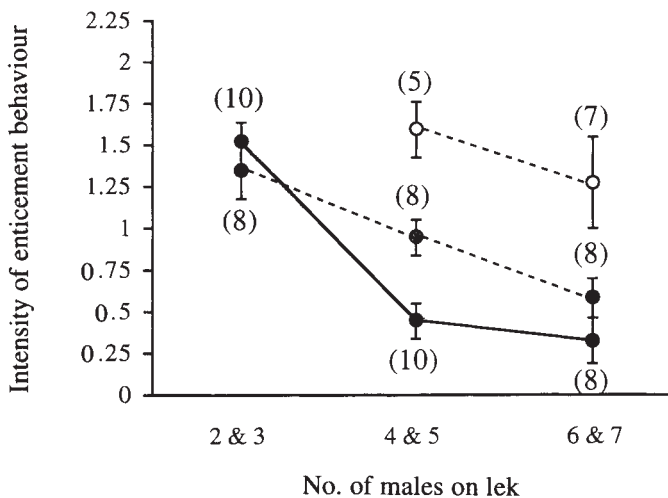


FIG. 4 Observed relationships between mean intensity of enticement behaviour towards flying males and current lek size across males of different rank. The two top-ranking males are represented by black circles, males with rank 3 and 4 are represented by hatched circles, and males of rank 5 and 6 are represented by white circles. The intensity of enticement behaviour to males differed significantly between males of different rank on lek size 4–5 (ANOVA; $F_{2,19} = 13.72$, $P < 0.0001$) and lek size 6–7 (ANOVA; $F_{2,19} = 8.82$, $P < 0.01$), but not at lek size 2–3 (ANOVA; $F_{2,17} = 3.35$, $P > 0.10$). There were no significant differences between the male groups in intensity of enticement behaviour towards flocks containing females at any lek size (ANOVAs; $F_{2,15} < 1.73$, $P > 0.10$). Male ruffs actively try to entice both females and males to leks by displaying the white underparts of their wings to flying conspecifics. For each enticement event, enticement displays were ranked on a four point scale. Males that did not entice scored as 0, males that flapped their wings while standing scored as 1, males that flapped while jumping up and down scored as 2, and males that hovered over the lek scored as 3. Data come from morning observations within 10 days of the annual population mating peak in 1993; 25 males from 5 different leks were included in the analysis, and the mean enticement intensity for all events present and respective lek size was calculated for each male. Values are means ($\pm$ s.e.) and numbers in parentheses refer to number of males. Males and lek sizes were grouped to increase the power of the tests, and all observations with females already present on the lek were excluded from the analyses.

tion of some degree of communal displaying, the tight clustering and large aggregations found in most lekking species probably represent the selfish interests of low-ranking males alone.

We suggest that the interaction between lek size and mating skew may be an important factor in promoting aggregated displays, while at the same time setting the upper limit to lek size. In combination with female preferences for traditional lek sites in areas of suitable breeding habitat, this could be an important component in the explanation of the evolution of lekking. Further empirical data suggest that the patterns we have observed in the ruff are common to several other species of lekking birds and mammals (F.W. and I.P.F.O., manuscript in preparation), and the suggested mechanism for the evolution of aggregated displays, therefore, appears to be general. Future studies should focus on both degree of male mating skew and copulation rates, when discussing how individual males should distribute themselves. □

Received 24 June; accepted 24 November 1994.

- Bradbury, J. W. in *Natural Selection and Social Behaviour* (eds Alexander, R. D. & Tinkle, D. W.) 138–169 (Chiron, New York, 1981).
- Bradbury, J. W. & Gibson, R. M. in *Mate Choice* (ed. Bateson, P.) 109–138 (Cambridge Univ. Press, 1983).
- Bradbury, J. et al. *Anim. Behav.* **34**, 1694–1709 (1986).
- Beehler, B. M. & Foster, M. S. *Am. Nat.* **131**, 203–219 (1988).
- Alatalo, R. V. et al. *Behav. Ecol.* **3**, 53–59 (1992).
- Clutton-Brock, T. H. et al. *Behav. Ecol.* **3**, 234–243 (1992).

- Höglund, J. et al. *Behav. Ecol. Sociobiol.* **32**, 31–39 (1993).
- Stillman, R. A. et al. *Behav. Ecol.* **4**, 1–6 (1993).
- Davies, N. B. in *Behavioural Ecology* (eds Krebs, J. R. & Davies, N. B.) 263–294 (Blackwell, Oxford, 1991).
- Wiley, R. H. *Adv. Study Behav.* **20**, 202–291 (1991).
- Milinski, M. & Parker, G. A. in *Behavioural Ecology* (eds Krebs, J. R. & Davies, N. B.) 137–168 (Blackwell, Oxford, 1991).
- Hogan-Waenburg, A. J. *Ardea* **54**, 8–229 (1966).
- Lank, D. B. & Smith, C. M. *Behav. Ecol. Sociobiol.* **30**, 323–329 (1992).
- van Rhijn, J. G. *The Ruff* (Poyser, London, 1991).
- Keller, L. & Vargo, E. L. in *Queen Number and Sociality in Insects* (ed. Keller, L.), 16–44 (Oxford Univ. Press, 1993).
- Reeve, H. K. & Ratnicks, F. L. W. in *Queen Number and Sociality in Insects* (ed. Keller, L.), 45–85 (Oxford Univ. Press, 1993).

ACKNOWLEDGEMENTS. We thank S. Albon, J. Höglund, A. P. Möller and S. Ulfstrand for helpful comments on an earlier draft. S. Albon also kindly helped with the statistical analyses. This paper was written while F.W. was visiting the Institute of Zoology, London, on a Royal Swedish Academy of Sciences grant. The field work was funded by RSAS and WWF grants to F.W. and I.P.F.O. was supported by a NERC Postdoctoral Fellowship.

Reduced hippocampal LTP and spatial learning in mice lacking NMDA receptor $\epsilon 1$ subunit

Kenji Sakimura*, Tatsuya Kutsuwada*, Isao Ito†, Toshiya Manabe‡, Chitoshi Takayama§, Etsuko Kushiya*, Takeshi Yagi¶, Shinichi Aizawa¶, Yoshiro Inoue§, Hiroyuki Sugiyama† & Masayoshi Mishina*^{##}

* Department of Neuropharmacology, Brain Research Institute, Niigata University, Niigata 951, Japan

† Department of Biology, Faculty of Science, Kyushu University, Fukuoka 812, Japan

‡ Department of Neurophysiology, Institute for Brain Research, Faculty of Medicine, University of Tokyo, Tokyo 113, Japan

§ Department of Anatomy, Hokkaido University School of Medicine, Sapporo 060, Japan

¶ Laboratory of Molecular Oncology, Tsukuba Life Science Center, Institute of Physical and Chemical Research, Ibaraki 305, Japan

^{##} Department of Pharmacology, Faculty of Medicine, University of Tokyo, Tokyo 113, Japan

THE NMDA (*N*-methyl-D-aspartate) receptor channel is important for synaptic plasticity, which is thought to underlie learning, memory and development^{1,2}. The NMDA receptor channel is formed by at least two members of the glutamate receptor (GluR) channel subunit families, the GluR ϵ (NR2) and GluR ζ (NR1) subunit families^{3–8}. The four ϵ subunits are distinct in distribution, properties and regulation^{5–14}. On the basis of the Mg^{2+} sensitivity and expression patterns, we have proposed that the $\epsilon 1$ (NR2A) and $\epsilon 2$ (NR2B) subunits play a role in synaptic plasticity^{6,14}. Here we show that targeted disruption of the mouse $\epsilon 1$ subunit gene resulted in significant reduction of the NMDA receptor channel current and long-term potentiation at the hippocampal CA1 synapses. The mutant mice also showed a moderate deficiency in spatial learning. These results support the notion that the NMDA receptor channel-dependent synaptic plasticity is the cellular basis of certain forms of learning.

To disrupt the GluR $\epsilon 1$ locus, the targeting vector shown in Fig. 1a was transfected into murine TT2 embryonic stem (ES) cells¹⁵. Two targeted clones transmitted the mutation through the germ line, and their progenies were analysed by Southern blot hybridization (Fig. 1b). Homozygous GluR $\epsilon 1$ mutant mice

* Present addresses: Laboratory of Neurobiology and Behavioral Genetics, National Institute for Physiological Sciences, Okazaki 444, Japan (T.Y.) and Laboratory of Morphogenesis, Institute of Molecular Embryology and Genetics, Kumamoto University School of Medicine, Kumamoto 860, Japan (S.A.).

^{##} To whom correspondence should be addressed at the Department of Pharmacology, Faculty of Medicine, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113, Japan.

developed and mated normally, in contrast to the lethal effect of disruption of the *GluR ζ 1* (NR1) gene¹⁶. The mutant mice seemed to be 'jumpy'. *In situ* hybridization analysis of the sagittal brain sections with oligonucleotide probes corresponding to the disrupted (probe ϵ 1C) and intact (probe ϵ 1A) regions showed that normal *GluR ϵ 1*-RNA transcripts were absent in mutant mice (Fig. 1c). Expression of messenger RNAs corresponding to the other NMDA receptor channel subunits were not appreciably affected by the *GluR ϵ 1* disruption. Western blot analysis of forebrain proteins showed that the mutant mice (-/-) contained no detectable *GluR ϵ 1* protein (M_r 175K), and the content of the heterozygous mice (+/-) was approximately half that of the wild-type mice (+/+), whereas the contents of *GluR ϵ 2* protein (180K) and the large myelin-associated glycoprotein (LMAG) were not appreciably affected by the mutation (Fig. 1d).

Histological examination of various brain regions where the ϵ 1 subunit mRNA was abundantly expressed, such as the hippocampal formation, cerebral cortex, olfactory bulb, ventral posterior nucleus of the thalamus, spinal tract nucleus of trigeminal nerve and inferior olive^{5,7,9,10} showed no obvious morphological abnormalities in the *GluR ϵ 1*-mutant mice (Fig. 2). The ϵ 1 subunit mRNA is expressed only postnatally⁹ and thus may exert little effect on development.

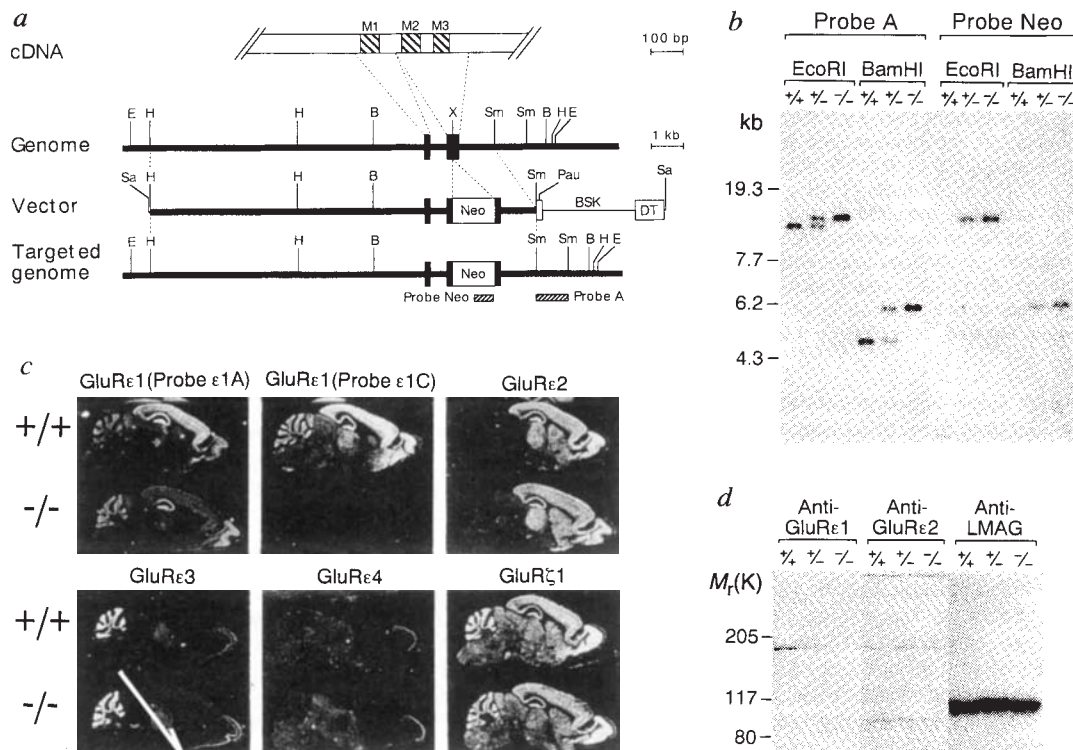
Efficacy of excitatory transmission in the hippocampal CA1 pyramidal neurons, estimated by input-output relationships of field excitatory postsynaptic potentials (e.p.s.ps), was indistinguishable between the wild-type and mutant slices, and the

time courses of the field potential responses were also similar (Fig. 3a). No significant difference in the extent of paired-pulse facilitation was found between the wild-type (1.75 ± 0.04 , mean $\pm$ s.e.m., $n=10$) and mutant (1.67 ± 0.03 , $n=10$) slices. But NMDA receptor channel currents, expressed as the ratio to 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX)-sensitive non-NMDA receptor channel currents, were significantly reduced in the mutant slices ($19 \pm 3\%$, $n=5$) compared with the wild-type slices ($43 \pm 6\%$, $n=5$) (t -test, $P<0.01$) (Fig. 3b). On the other hand, there was no significant difference in the current-voltage relationship of the NMDA receptor channel currents (Fig. 3c), and the reversal potential in the mutant slices (2 ± 1 mV, $n=7$) was similar to that in the wild-type slices (3 ± 2 mV, $n=6$).

Tetanic stimulation caused a long-lasting increase in the synaptic strength in the wild-type slices; the relative e.p.s.p. slope 60 min after tetanic stimulation being $187 \pm 8\%$ of the average slope before stimulation ($n=20$) (Fig. 3d). But the extent of the long-lasting potentiation was significantly decreased in the mutant slices ($126 \pm 6\%$, $n=17$) ($P<0.01$). In mutant mice defective in the γ isoform of protein kinase C (PKC γ), an apparent defect in hippocampal long-term potentiation (LTP) was fully removed when low-frequency stimulation was applied before tetanic stimulation¹⁷. The extent of potentiation induced by tetanic stimulation after low-frequency stimulation was still significantly smaller in the mutant mice ($162 \pm 15\%$, $n=8$) than in the wild-type mice ($209 \pm 15\%$, $n=8$) ($P<0.05$) (Fig. 3e).

As a control for the Morris water-maze task¹⁸, mice were first tested on a visible-platform task. Although the mutant mice took

FIG. 1 Targeted disruption of the mouse NMDA receptor channel ϵ 1 subunit gene. **a**, Schematic representations of *GluR ϵ 1* complementary DNA, genomic DNA, targeting vector and disrupted gene. M1-3, putative transmembrane segments; Neo, neomycin resistant gene; Pau, mRNA-destabilizing and transcription-pausing signals²⁵; BSK, plasmid pBluescript; DT, diphtheria toxin gene; B, *Bam*HI; E, *Eco*RI; H, *Hind*III; Sa, *Sac*I; Sm, *Sma*I; X, *Xcm*I. **b**, Southern blot analysis of *Eco*RI- or *Bam*HI-digested genomic DNA. **c**, *In situ* hybridization analysis of parasagittal brain sections. **d**, Western blot analysis of forebrain proteins with anti-*GluR ϵ 1*, anti-*GluR ϵ 2* and anti-LMAG antibodies. METHODS. The targeting vector pTVGR ϵ 1 was constructed by inserting the 1.3-kilobase (kb) blunted *Eco*RI-*Bam*HI fragment from pGK2Neo containing the neomycin phosphotransferase gene driven by the phosphoglycerate kinase promoter²⁵ into the blunted *Xcm*I site of 10-kb *GluR ϵ 1* genomic DNA isolated from a C57BL/6 genomic library (provided by A. Shimizu) and by ligating pPauDT1 for negative selection, a derivative of pHFZprNeoDT²⁵, in which the multi-cloning sequence between the *Cla*I and *Sac*I sites was removed and a synthetic adapter carrying *Sac*I, *Hind*III and *Sma*I sites replaced the 16.5-kb *Not*I-*Xho*I segment. TT2 (C57BL/6 $\times$ CBA F₁ hybrid) ES cells¹⁵ were transfected by *Not*I-digested pTVGR ϵ 1, and five targeted clones were identified by G418 selection, polymerase chain reaction and Southern blot hybridization, the frequency being 0.4% of G418-resistant clones. Chimaeras produced by injection of these cells into 8-cell embryos of



ICR mice were mated with C57BL/6 mice and clones H10 and N6 transmitted the mutation through the germ line. The contribution of the ϵ 1 subunit is larger in adult animals⁹, thus all analyses were done with mature mice (9–10 weeks old), using wild-type littermates as controls. Essentially the same results were obtained with two lines and data from H10 are shown. *In situ* hybridization and western blot analyses were done as described^{9,10,26}. Oligonucleotide probe ϵ 1C is complementary to the nucleotide residues 1,795 to 1,839 of the ϵ 1 subunit cDNA⁵. Anti-*GluR ϵ 1* and anti-*GluR ϵ 2* antibodies were prepared using fusion proteins of a carboxy-terminal portion of the ϵ 1 (amino-acid residues 1,335 to 1,414)⁵ or ϵ 2 subunit (residues 1,353 to 1,432)⁶ with glutathione-S-transferase²⁶. Anti-LMAG antibody provided by T. Inuzuka.

FIG. 2 Nissl-stained brain sections of the wild-type (+/+), mutant (-/-) mice. *a, b*, Parasagittal sections of the whole brain. *c, e*, Coronal sections of the parietal lobe. *d, f*, Coronal sections of the hippocampal formation. Mice were fixed by transcardial perfusion with 4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4. Cryostat-prepared 20- μ m sections were stained with toluidine blue. BS, Brainstem; CA1 and CA3 are fields of the Ammon's horn; Cb, cerebellum; Cx, cerebral cortex; DG, dentate gyrus; Ha, habenular nucleus; Hi, hippocampal formation; OB, olfactory bulb; Th, thalamus; I-VI, cortical laminae I-VI.

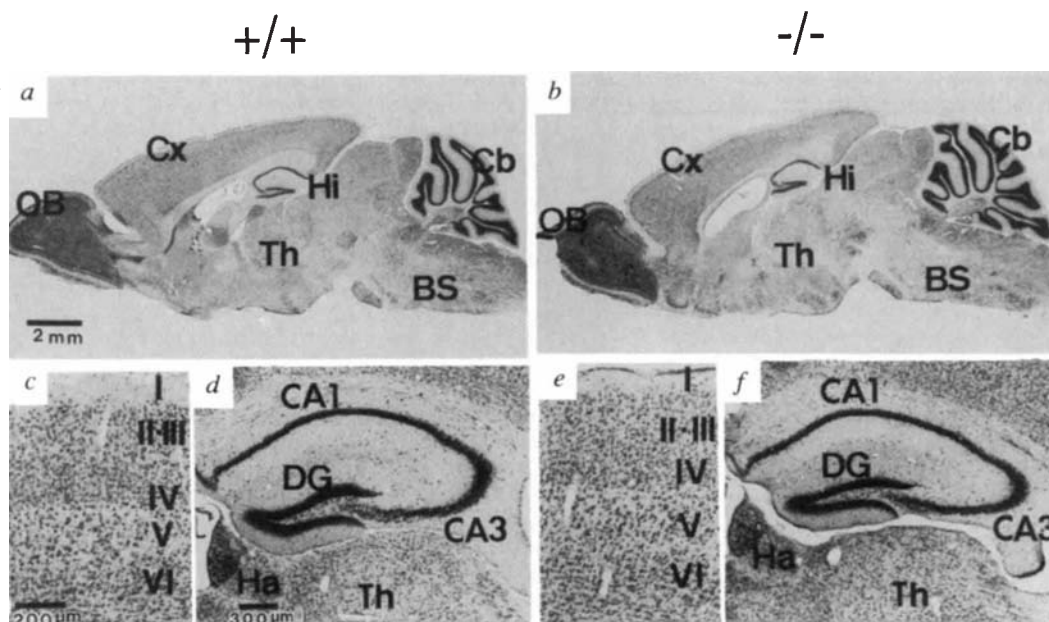
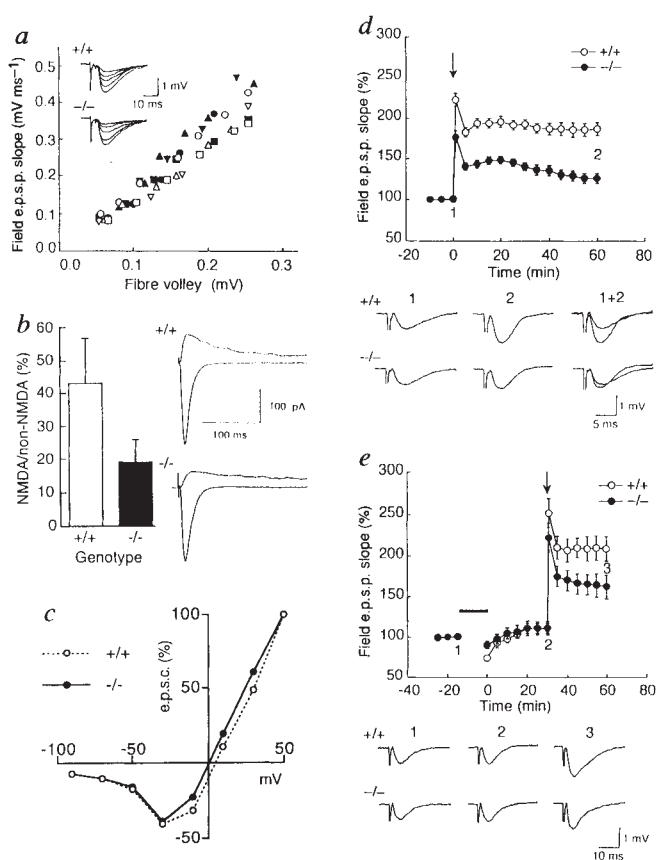


FIG. 3 Synaptic responses in hippocampal CA1 pyramidal cells of the wild-type (open symbols) and mutant (filled symbols) mice. *a*, Initial slopes of field e.p.s.ps as a function of presynaptic fibre volley amplitudes. Each symbol represents a set of experiments from a single slice. Inset traces show typical field e.p.s.ps obtained with various stimulus intensities. *b*, Ratios of peak NMDA receptor channel currents at +40 mV to peak synaptic currents at -90 mV, measured with the whole-cell voltage-clamp recording. The NMDA receptor-mediated current recorded at +40 mV was blocked by 25 μ M D-APV. Upper traces show synaptic currents at +40 mV in the presence of 20 μ M CNQX, and lower traces currents at -90 mV in the control solution. Ratios of the two currents in the wild-type and mutant slices measured with the intracellular recording under discontinuous voltage-clamp were $34 \pm 11\%$ and $12 \pm 3\%$ ($n=6$ each), respectively, at -40 mV holding potential. *c*, Representative current-voltage relationships of NMDA receptor channel currents measured in the presence of 20 μ M CNQX. The currents were normalized to the values at +50 mV e.p.s.c., excitatory post-synaptic current. *d, e*, LTP of the hippocampal CA1 field e.p.s.p. expressed as percentage of the mean before tetanic stimulation. The control stimulation intensity was $\sim 30\%$ of the maximal response and it was kept constant throughout the experiment. A single slice was used from one animal. *d*, Potentiation induced by tetanic stimulation (100 Hz for 1 s, 2 trains, 30-s interval, arrow) in 20 wild-type slices (open circles) and 17 mutant slices (filled circles). *e*, Effect of previous low-frequency stimulation (1 Hz for 900 s, thick bar) on potentiation induced by tetanic stimulation (100 Hz for 1 s, arrow); eight slices each from the wild-type (open circles) and mutant (filled circles) mice. Typical traces (average of three recordings) at the times indicated by Arabic numerals are shown below.

METHODS. The series of experiments included some performed blind (47%) and essentially the same results were obtained. Transverse hippocampal slices ($\sim 400 \mu$ m) were placed in a submerged recording chamber perfused continuously with artificial cerebrospinal fluid (ACSF) equilibrated with 95% $O_2/5\%$ CO_2 at $32 \pm 0.5^\circ C$ at a rate of 3–4 ml min^{-1} . Extracellular field e.p.s.ps were recorded with an electrode filled with 0.9% (NaCl) in CA1 stratum radiatum²⁷. A bipolar tungsten electrode was used to stimulate Schaffer collateral/commissural afferents $\sim 200 \mu$ m away from the recording electrode. Stimulus duration, 80 μ s. For measurements of paired-pulse facilitation, paired pulses were given at 40 ms intervals. ACSF contained 119 mM NaCl, 2.5 mM KCl, 2.5 mM $CaCl_2$, 1.3 mM $MgSO_4$, 1 mM NaH_2PO_4 , 26 mM $NaHCO_3$ and 10 mM glucose. E.p.s.cs were recorded from CA1 pyramidal cells with a patch electrode (3–6 M Ω) in the whole-cell voltage-clamp mode (Axopatch 1D) using the 'blind' technique^{28,29}. The pipette solution contained 122.5 mM caesium gluconate, 17.5 mM CsCl, 10 mM HEPES buffer,



0.2 mM EGTA, 8 mM NaCl, 2 mM $Mg-ATP$, 0.3 mM Na_3-GTP (pH 7.2, 290–300 mOsm). The values of the membrane potential were corrected for the liquid-junction potential at the electrode tip. The discontinuous voltage clamp (holding potential, -40 mV; sampling clock, 1–2 kHz) was used to record e.p.s.cs at $25 \pm 5^\circ C$ with an intracellular electrode (30–40 M Ω) filled with 3 M CsCl. Picrotoxin (50–100 μ M) was present in all experiments.

longer to reach the platform on the first block of trials, they subsequently performed as efficiently as the wild-type mice (Fig. 4a). Thus, the GluR ϵ 1-mutant and wild-type mice are able to perform the non-spatial learning water-maze task. In the hidden-platform task, the wild-type mice quickly learned to locate the hidden platform, whereas the mutant mice took significantly longer (Fig. 4b).

In the transfer test², the wild-type mice spent longer in the trained quadrant than in the opposite, right or left quadrant (Fig. 4c), and crossed the exact spot where the platform had been located more often than they did equivalent locations in other quadrants (Fig. 4d). The mutant mice also searched preferentially in the trained quadrant, but not as selectively as the wild-type mice; the difference in searching time between the trained and right quadrants was less significant. The searching time in the trained quadrant by the mutant mice was significantly shorter than by the wild-type mice ($P < 0.01$). Furthermore, the mutant mice failed to cross the trained site selectively; there was no significant difference in the number of times that they cross the trained site compared with the equivalent location in the right-hand quadrant. The mutant mice crossed the trained spot less often than the wild-type mice ($P < 0.01$). These results suggest that the GluR ϵ 1-mutant mice perform the spatial learning task less precisely than the wild-type mice.

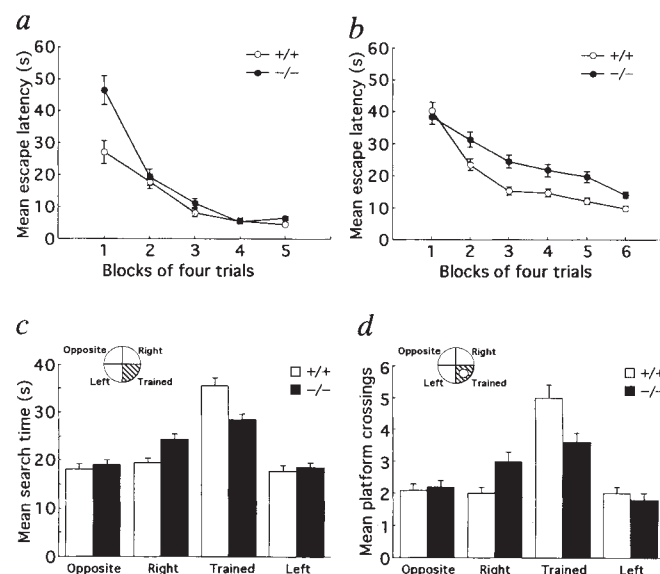
Through the analyses of expression patterns and functional properties of diverse NMDA receptor channel subunits identified by molecular cloning, we have proposed that of the four ϵ subunits, the ϵ 1 and ϵ 2 subunits are involved in synaptic plasticity^{6,14}. Here we have provided experimental evidence that the ϵ 1 subunit is a component of functional NMDA receptor channels in the hippocampus and is important for LTP of synaptic transmission and spatial learning. Because Ca^{2+} entry through NMDA receptor channels triggers the induction of LTP in the hippocampal CA1 region¹, the decrease in NMDA receptor channel activity should causally relate to the reduction of LTP in the GluR ϵ 1 mutant mice. The observation that hippocampal LTP is not completely abolished by the deprivation of the ϵ 1 subunit implies that the ϵ 2 subunit would also play a role

in the hippocampal LTP, as both the ϵ 1 and ϵ 2 subunit mRNAs are expressed in the adult hippocampus^{5,6,9,10} and share similar functional properties such as high sensitivity to Mg^{2+} blockage and modulation^{5,7,11,13}. The GluR ϵ 1-mutant mice are apparently jumpy and seem to have enhanced startle response, similar to mutant mice defective in α isoform of Ca^{2+} -calmodulin-dependent protein kinase II (α -CaMKII)¹⁹. It is possible that disturbance of LTP in amygdala for example, might be related to the behaviour¹⁹.

Synaptic plasticity, represented by LTP in the hippocampus, has long been thought to underlie learning and memory¹, but experimental evidence for the linkage is limited. The initial approach demonstrated that chronic intraventricular infusion of 2-amino-5-phosphonovalerate (APV), an antagonist of the NMDA receptor channel, impaired both hippocampal LTP and spatial learning in rat^{2,20}. We have shown by gene targeting that the GluR ϵ 1-mutant mice, showing reduced hippocampal LTP, have a moderate defect in spatial learning, which is quantitatively similar to results obtained with an intermediate dose of APV (ref. 20), supporting the notion that the NMDA receptor channel-dependent synaptic plasticity is the cellular basis of certain forms of learning. But it cannot be excluded that possible disturbance of synaptic transmission, undetected slight abnormalities of the brain organization or the jumpy behaviour of the GluR ϵ 1-mutant mice might affect the observed deficiency in spatial learning. Mice defective in α -CaMKII or the non-receptor tyrosine kinase *fyn*, showed a deficiency in hippocampal LTP and spatial learning, though the exact mechanisms were unknown^{19,21,22}. These effects may be indirect because in *fyn*-mutant mice, defects were noted in the arrangement of hippocampal neurons and myelination^{21,23}. PKC γ -mutant mice showed spatial learning capabilities in spite of an apparent defect in hippocampal LTP^{17,24}, and LTP could be fully induced in these mice when low-frequency stimulation was applied before tetanic stimulation. In the GluR ϵ 1-mutant mice, LTP was significantly reduced even after such stimulation. Our findings taken together are in line with the notion that synaptic plasticity is the cellular mechanism of certain forms of memory and learning. □

FIG. 4 Performance in the Morris water maze tasks by the wild-type (open symbols) and mutant (filled symbols) mice. **a**, Escape latency in the visible-platform test ($n = 16$, each). Analysis of variance (ANOVA) with repeated measures showed significant effects of genotypes ($F_{(1,30)} = 11.71$, $P < 0.01$) and trial blocks ($F_{(4,120)} = 78.74$, $P < 0.01$). The interaction between trial blocks and genotypes ($F_{(4,120)} = 7.24$, $P < 0.01$) was significant only at the first block ($F_{(1,30)} = 11.51$, $P < 0.01$). **b**, Escape latency in the hidden-platform test ($n = 40$, each). The effects of genotypes ($F_{(1,78)} = 12.03$, $P < 0.01$) and trial blocks ($F_{(5,390)} = 53.89$, $P < 0.01$) were significant, and there was no interaction. There was no significant difference in body weight and swimming speed between the wild-type (21.10 ± 0.46 g and 24.27 ± 0.33 cm s⁻¹, respectively) and mutant (20.96 ± 0.40 g and 24.22 ± 0.31 cm s⁻¹, respectively) mice. **c**, **d**, Performance in the transfer test ($n = 40$, each), on the day after hidden-platform training. **c**, Quadrant search time. The wild-type mice spent more time in the trained than in the other three quadrants; ANOVA, $F_{(2,156)} = 67.15$, $P < 0.01$; Tukey analysis, trained quadrant > other quadrants, $P < 0.01$. Mutant mice also spent more time in the trained quadrant; $F_{(2,156)} = 30.57$, $P < 0.01$; Tukey analysis, trained > opposite and left, $P < 0.01$, trained > right, $P < 0.05$. **d**, Number of crossings. The wild-type mice crossed the trained site more often than equivalent sites in other quadrants; $F_{(3,156)} = 26.12$, $P < 0.01$; Tukey analysis, trained > others, $P < 0.01$. Mutant mice failed to cross the trained site selectively; $F_{(3,156)} = 9.84$, $P < 0.01$; Tukey analysis, trained = right, trained > opposite and left, $P < 0.01$.

METHODS. The visible- and hidden-platform tests were done as described^{21,30} with slight modifications. The series of experiments included 63% performed blind and essentially the same results were obtained. The training apparatus was a circular white-painted steel pool, 120 cm in diameter. The water was made opaque by adding skimmed milk and was kept at 18.5 ± 0.5 °C. Before the test, mice were first trained to swim and climb on the platform. In the visible-platform test,



mice were given one block of four trials on the first day and two blocks of trials on the second and third days. In the hidden-platform test, mice were given one trial of the visible-platform task as pretraining one day before the test and were given one block of four trials per day for six consecutive days. Swimming paths were tracked with a camera and stored in a computer (TARGET/2 system, Neuroscience Inc., Tokyo).

Received 27 July; accepted 22 November 1994.

- Bliss, T. V. P. & Collingridge, G. L. *Nature* **361**, 31–39 (1993).
- Morris, R. G. M., Anderson, E., Lynch, G. S. & Baudry, M. *Nature* **319**, 774–776 (1986).
- Moriyoshi, K. et al. *Nature* **354**, 31–37 (1991).
- Yamazaki, M. et al. *FEBS Lett.* **300**, 39–45 (1992).
- Meguro, H. et al. *Nature* **357**, 70–74 (1992).
- Kutsuwada, T. et al. *Nature* **358**, 36–41 (1992).
- Monyer, H. et al. *Science* **256**, 1217–1221 (1992).
- Ikeda, K. et al. *FEBS Lett.* **313**, 34–38 (1992).
- Watanabe, M., Inoue, Y., Sakimura, K. & Mishina, M. *NeuroReport* **3**, 1138–1140 (1992).
- Watanabe, M., Inoue, Y., Sakimura, K. & Mishina, M. *J. comp. Neurol.* **338**, 337–390 (1993).
- Mori, H., Yamakura, T., Masaki, H. & Mishina, M. *NeuroReport* **4**, 519–522 (1993).
- Ishii, T. et al. *J. biol. Chem.* **268**, 2836–2843 (1993).
- Monyer, H., Burnashev, N., Laurie, D. J., Sakmann, B. & Seeburg, P. H. *Neuron* **12**, 529–540 (1994).
- Mishina, M. et al. *Ann. N.Y. Acad. Sci.* **707**, 136–152 (1993).
- Yagi, T. et al. *Analyt. Biochem.* **214**, 70–76 (1993).
- Li, Y. et al. *Cell* **76**, 427–437 (1994).
- Abeliovich, A. et al. *Cell* **75**, 1253–1262 (1993).
- Morris, R. G. M. *Learn. Motiv.* **12**, 239–260 (1981).
- Silva, A. J., Stevens, C. F., Tonegawa, S. & Wang, Y. *Science* **257**, 201–206 (1992).
- Davis, S., Butcher, S. P. & Morris, R. G. M. *J. Neurosci.* **12**, 21–34 (1992).
- Silva, A. J., Paylor, R., Wehner, J. M. & Tonegawa, S. *Science* **257**, 206–211 (1992).
- Grant, S. G. N. et al. *Science* **258**, 1903–1910 (1992).
- Umemori, H., Sato, S., Yagi, T., Aizawa, S. & Yamamoto, T. *Nature* **367**, 572–576 (1994).
- Abeliovich, A. et al. *Cell* **75**, 1263–1271 (1993).
- Yagi, T. et al. *Analyt. Biochem.* **214**, 77–86 (1993).
- Araki, K. et al. *Biochem. biophys. Res. Commun.* **197**, 1267–1276 (1993).
- Ito, I. & Sugiyama, H. *NeuroReport* **2**, 333–336 (1991).
- Blanton, M. G., Lo Turco, J. J. & Kriegstein, A. R. *J. Neurosci. Meth.* **30**, 203–210 (1989).
- Coleman, P. A. & Miller, R. F. *J. Neurophysiol.* **61**, 218–230 (1989).
- Zhang, Y., Saito, H. & Nishiyama, N. *Brain Res.* **658**, 127–134 (1994).

ACKNOWLEDGEMENTS. The first two authors contributed equally to the work. We thank R. Natsume for assistance in animal care and DNA analyses, N. Takeda for help in the preparation of chimeric mice, H. Saito and N. Nishiyama for advice on the water maze task, M. Nakazawa for advice on statistical analyses, A. Shimizu for a mouse genomic DNA library, T. Inuzuka for anti-LMAG antibody, and H. Niki and T. Takahashi for critically reading the manuscript. This investigation was supported by the Ministry of Education, Science and Culture of Japan, the Uehara Memorial Foundation, the Ministry of Health and Welfare of Japan, the Naito Foundation and the Yamanouchi Foundation.

Subthreshold synaptic Ca^{2+} signalling in fine dendrites and spines of cerebellar Purkinje neurons

Jens Eilers, George J. Augustine* & Arthur Konnerth†

I. Physiologisches Institut der Universität des Saarlandes, D-66421 Homburg, Germany

* Department of Neurobiology, Duke University Medical Center, Durham, North Carolina 27710, USA

THE conventional view of synaptic integration is that it results from the simple summation of electrical signals produced by each active synapse innervating a given neuron¹. However, because synaptic action can go beyond the production of postsynaptic electrical signals, to include intracellular biochemical events such as the generation of second messengers, it is possible that synaptic integration could occur at another level². We have considered this possibility by examining changes in the dendritic concentration of the second messenger, calcium, resulting from subthreshold excitatory synaptic activity in cerebellar Purkinje neurons. We report here clear evidence that such non-electrical synaptic integration occurs and that it takes place in restricted dendritic compartments consisting of spines and adjacent fine dendrites.

By combining whole-cell patch-clamp recordings with confocal laser scanning microscopy (Fig. 1a) we could measure postsynaptic Ca^{2+} signals from the finest, terminal spiny dendrites

of single Purkinje neurons of cerebellar slices (Figs 1b and 2). Activation of excitatory, parallel fibre (PF) synapses produced multiple types of dendritic Ca^{2+} changes. Single-shock activation of a sufficient number of PF synapses to produce a subthreshold excitatory postsynaptic potential (e.p.s.p.), as in Fig. 1c, produced highly localized but clearly visible rises in postsynaptic Ca^{2+} activity (Fig. 1d). Such Ca^{2+} signals were readily found in all Purkinje neurons that were tested ($n=37$). These rises in postsynaptic Ca^{2+} concentration occurred at levels of synaptic activity that were below the threshold for regenerative electrical activity, such as action or plateau potentials³. Stimulation of a larger number of PFs, a number sufficient to reach the threshold for producing a propagated action potential (the so-called 'simple spike'), also produced a rise in postsynaptic Ca^{2+} concentration that was quite similar to that produced by subthreshold e.p.s.ps (Fig. 1e,f). These Ca^{2+} signals were confined to the same compartment of terminal spiny dendrites, but resulted in larger rises in postsynaptic Ca^{2+} concentration within the activated compartment. Still stronger and repetitive stimulation caused Ca^{2+} changes that extended beyond spiny dendrites and included tertiary and secondary dendrites, confirming previous observations^{4,5}. Even larger e.p.s.ps, sufficient to activate dendritic Ca^{2+} spikes^{5,9}, produced very large Ca^{2+} signals throughout the Purkinje neuron⁵ (data not shown).

The postsynaptic Ca^{2+} signals resulting from subthreshold PF synaptic activity were not restricted to individual dendrites or spines, but instead appeared in a restricted area consisting of both (Fig. 2). In the example shown in Fig. 2, a single subthreshold PF-e.p.s.p. produced rises in Ca^{2+} concentration in spines (trace 1) as well as in the larger dendritic structure to which these spines were attached (trace 3). This was a general finding, with the dendritic Ca^{2+} signals usually found to terminate abruptly within the dendrites, over distances of just a few micrometres. The spatial extent of these postsynaptic Ca^{2+} signals may be limited by the failure of the electrical signals to propagate beyond the points at which the dendrites branch within the dendritic arborization. The fact that these Ca^{2+} signals are not restricted to the dendritic spines, where the PF terminals contact the Purkinje neurons¹⁰, contrasts with the case of hippocampal CA3 pyramidal neurons, where NMDA receptor-mediated postsynaptic Ca^{2+} influx is localized to single dendritic spines¹¹. However, the spread of Ca^{2+} signals into both spines and dendrites is not an artefact resulting from optical blurring of signals emanating from only one of these structures, because extracellular regions immediately adjacent to dendritic spines showed no significant changes in fluorescence during synaptic activity (Fig. 2, trace 2). Further, the time courses of the postsynaptic Ca^{2+} signals produced by PF-e.p.s.ps appeared similar in both dendrites and their spines (Fig. 2). There was no detectable lag in the rise of the Ca^{2+} signals in these two compartments within the 33 ms time resolution of our imaging method. Thus, the presence of Ca^{2+} signals both in spines and in adjacent dendrites results from the fact that Ca^{2+} rises in both compartments. These results indicate that the subthreshold postsynaptic Ca^{2+} signals of Purkinje neurons are not restricted to dendritic spines and, in fact, the spines may not even be the initial source of the Ca^{2+} influx that produces these signals.

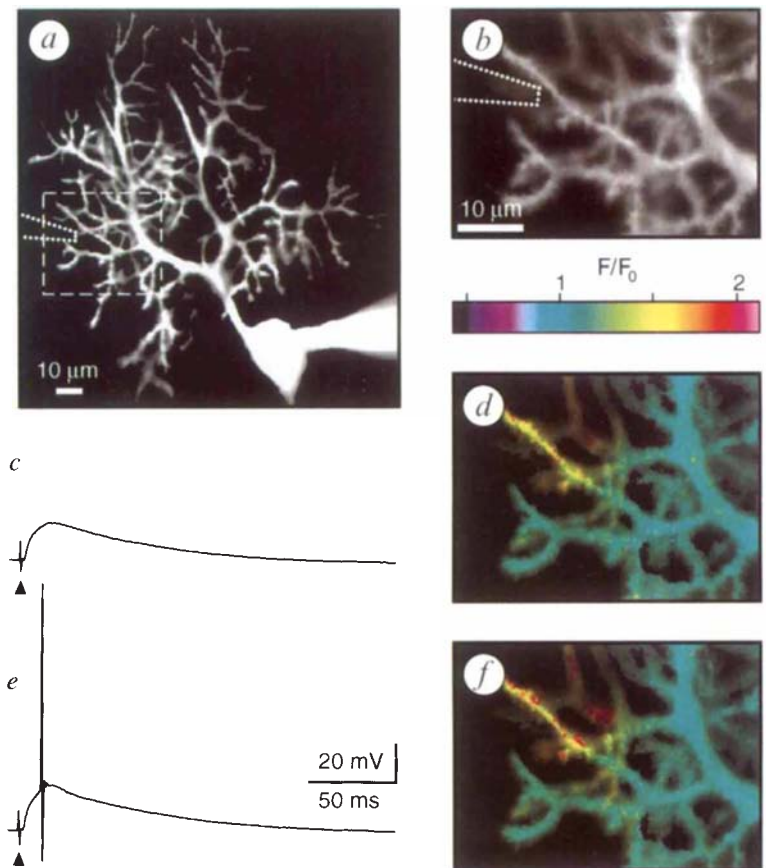
These Ca^{2+} signals were quite transient within both compartments, declining to half their maximum amplitude in less than 2 s (Fig. 2). This is an upper estimate of the time course of the Ca^{2+} signals, which could be even more rapid^{3,5} because of the Ca^{2+} -buffering properties of the fluorescent indicator^{12,13}. This brief time course also contrasts with the case of dendritic spines in hippocampal CA3 pyramidal neurons¹¹ but is much like the case for NMDA receptor-mediated postsynaptic Ca^{2+} signals in cerebral cortex¹⁴ and in hippocampal CA1 neurons¹⁵.

We next examined the sources of Ca^{2+} that produce the postsynaptic signals. These signals require activation of postsynaptic α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptors because a blocker of these receptors, 6-cy-

† To whom correspondence should be addressed.

FIG. 1 Synaptic Ca^{2+} signalling in terminal spiny dendrites of cerebellar Purkinje neurons. *a*, Confocal image of a patch-clamped Purkinje neuron containing the fluorescent calcium indicator dye, calcium green-1. The dendritic region delimited by the dashed box is shown at a higher magnification in *b*. The dotted lines in *a* and *b* indicate the position of an extracellular stimulation pipette. *c*, E.p.s.p. evoked by a single stimulus (100 μs , 4.0 V) applied to the parallel fibres. The arrowheads indicate the time of synaptic stimulation. $V_m = -70$ mV. *d*, Pseudocolour image of the relative increase in Ca^{2+} -dependent fluorescence (ratio between the peak fluorescence and the fluorescence before the stimulation, F/F_0) produced by the e.p.s.p. shown in *c*. *e*, Action potential caused by a suprathreshold e.p.s.p. that was evoked by a stimulus slightly stronger (100 μs , 4.3 V) than that used in *c*. *f*, Pseudocolour image of the Ca^{2+} signal associated with the suprathreshold e.p.s.p. shown in *e*.

METHODS. Experiments were done on 300- μm thick parasagittal cerebellar slices prepared from 18–22-day-old rats as previously described²⁸. Slices were perfused with a saline (20–23 °C) containing (in mM): 125 NaCl, 2.5 KCl, 2 CaCl_2 , 1 MgCl_2 , 1.25 NaH_2PO_4 , 26 NaHCO_3 , 20 glucose, 0.02 bicuculline (Sigma), bubbled with 95% O_2 and 5% CO_2 . The pipette solution contained (in mM): 132 KCl, 30 HEPES, 8 NaCl, 4 Mg-ATP, 0.4 $\text{Na}_3\text{-GTP}$, 0.5 calcium green-1 (Molecular Probes); pH 7.3. A confocal laser microscope (Noran, 'Odyssey', on Zeiss 'Axioskop', $\times 63$ water immersion objective, NA 0.9) was used to acquire fluorescence images at 30 Hz²⁹. The images in Figs 1*a*, *b* and 2 represent averages of 256 consecutive frames whereas those in Figs 1*d*, *f* and 4*B* are averages of 16.



ano-7-nitroquinoxaline-2,3-dione (CNQX, 10 μM , $n=7$) completely and reversibly blocked these postsynaptic Ca^{2+} signals and their associated synaptic currents (Fig. 3*a*). In contrast, these signals were not eliminated by treatments, such as bath application of (RS)- α -methyl-4-carboxyphenylglycine (MCPG, 500 μM , $n=4$, data not shown), that prevent metabotropic glutamate receptor (mGluR)-mediated Ca^{2+} signalling in other cell types¹⁶. Therefore, these Ca^{2+} signals could be a direct consequence of Ca^{2+} influx through the postsynaptic AMPA receptor^{13,17} or could be an indirect consequence of the e.p.s.p. depolarizing the postsynaptic membrane potential sufficiently to open the voltage-gated Ca^{2+} channels present in spiny dendrites^{18,19}. To distinguish between these possibilities, we asked whether keeping the postsynaptic membrane potential more constant prevented Ca^{2+} entry. We found that voltage-clamping the postsynaptic membrane potential did indeed reduce the amplitude of the Ca^{2+} signals (Fig. 3*b*), even though the dendritic membrane potential was not ideally clamped²⁰. Further, hyperpolarizing the membrane potential reduced the amplitude of the postsynaptic Ca^{2+} signals (Fig. 3*c*) while increasing the amplitude of the postsynaptic currents (Fig. 3*d*) ($n=4$). Thus, we conclude that the postsynaptic Ca^{2+} signals are due to AMPA receptors locally depolarizing the postsynaptic membrane potential to open voltage-gated Ca^{2+} channels in the region of dendritic membrane in the immediate vicinity of activated PFs. Interestingly, the postsynaptic Ca^{2+} signals have no regenerative electrical correlate that could be propagated down the Purkinje cell axon and communicated to its downstream target neurons²¹.

To evaluate the integrative properties of these postsynaptic Ca^{2+} signals, we stimulated different numbers of PF synapses by changing the intensity of the extracellular stimulation pulse²². Activating a small number of PFs produced no postsynaptic Ca^{2+} signal, even though these active PFs produced clear e.p.s.ps (Fig. 4*Ba*). However, increasing the number of active PFs

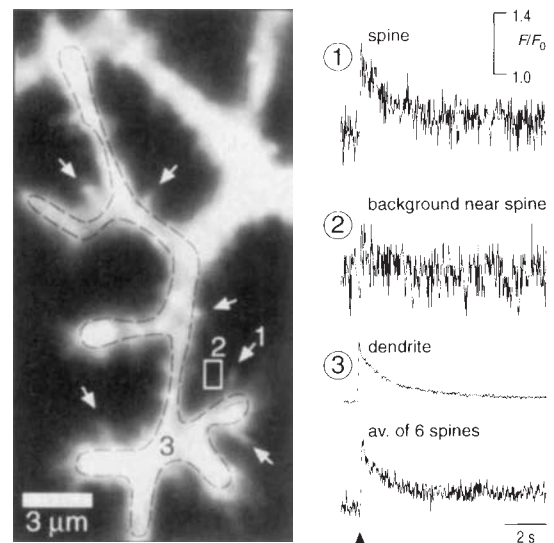


FIG. 2 Time course of subthreshold synaptic Ca^{2+} signals in spiny dendrites. The traces on the right represent Ca^{2+} signals measured at different sites during a single subthreshold parallel fibre-mediated e.p.s.p. with an amplitude of 9 mV. The upper trace (1) shows the time course of the Ca^{2+} signal measured in the spine marked with 1 on the left image. The corresponding background signal (trace 2) was recorded in the immediate vicinity of the spine (rectangle marked with 2). Trace 3 represents the change in fluorescence recorded within the integrated dendritic region delimited by the dashed line (3). The bottom trace was obtained by averaging Ca^{2+} signals recorded from the six individual spines marked by arrows. Spine fluorescence was measured within rectangles just covering each spine.

increased the size of the e.p.s.p. (Fig. 4A) and produced disproportionate increases in postsynaptic Ca^{2+} (Fig. 4Bb-d). These increases resulted from changes in the magnitude of the rise in postsynaptic Ca^{2+} concentration within the dendritic region, rather than increases in the dendritic area over which these Ca^{2+} rises occur (Fig. 4B). Quantification of the dependence of PF synaptic transmission and PF-mediated Ca^{2+} signals on stimulus strength emphasizes the nonlinear relationship between e.p.s.ps and Ca^{2+} signals (Fig. 4C). Based on this relationship, as well as the known size of e.p.s.ps produced by single PFs²³, we conclude that about 20 to 30 PFs need to be activated within the local dendritic area to produce a rise in postsynaptic Ca^{2+} concentration. Thus, the postsynaptic Ca^{2+} signals require summation of PF synaptic activity. This summation leads to a form of synaptic integration that is based on postsynaptic Ca^{2+} signalling rather than on conventional electrical signals.

In conclusion, our data reveal that large changes in postsynaptic Ca^{2+} concentration resulting from single subthreshold PF-e.p.s.ps occur in a strictly confined compartment consisting of a small number of branchlets of terminal spiny dendrites. This resultant postsynaptic Ca^{2+} signal has no direct electrical correlate and is therefore decidedly non-electrical in impact. Possible consequences of these local Ca^{2+} signals include downregulation of AMPA receptors^{24,25}, to modulate or even induce long-term depression at the PF to Purkinje cell synapse, and upregulation of GABA receptors to yield 'rebound potentiation' at inhibitory synapses²⁶. Thus, our results extend the notion of local hot spots of dendritic Ca^{2+} electrogenesis²⁷ by showing that the finest terminal spiny dendrites of the Purkinje neuron apparently form highly restricted functional units within which information about the number of active PF synapses is acquired and stored. □

FIG. 3 Ca^{2+} sources contributing to postsynaptic Ca^{2+} signals. **a**, The AMPA-receptor antagonist CNQX (10 μM) reversibly blocks the parallel-fibre mediated e.p.s.c. and the associated Ca^{2+} signal ($F_{\text{e.p.s.c.}}$). The pipette solution used in this particular experiment contained 500 μM fluo-3 (Molecular Probes) instead of 500 μM calcium green-1 that was used in all other experiments. $V_h = -60$ mV. **b**, Superposition of two Ca^{2+} signals associated with an e.p.s.c. ($F_{\text{e.p.s.c.}}$) and an e.p.s.p. ($F_{\text{e.p.s.p.}}$), recorded in the same Purkinje neuron by applying a similar stimulus to parallel fibres in current- and voltage-clamp conditions, respectively (see inset). Both recordings were at a membrane potential of -70 mV. **c**, Superposition of subthreshold synaptic Ca^{2+} signals recorded at the indicated holding potentials. **d**, The voltage dependence of e.p.s.c. (empty symbols) and Ca^{2+} signal (filled symbols) amplitudes.

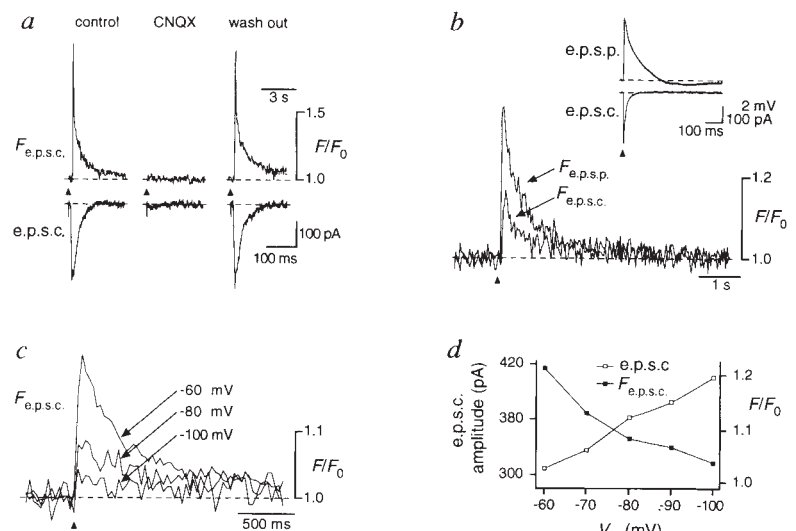
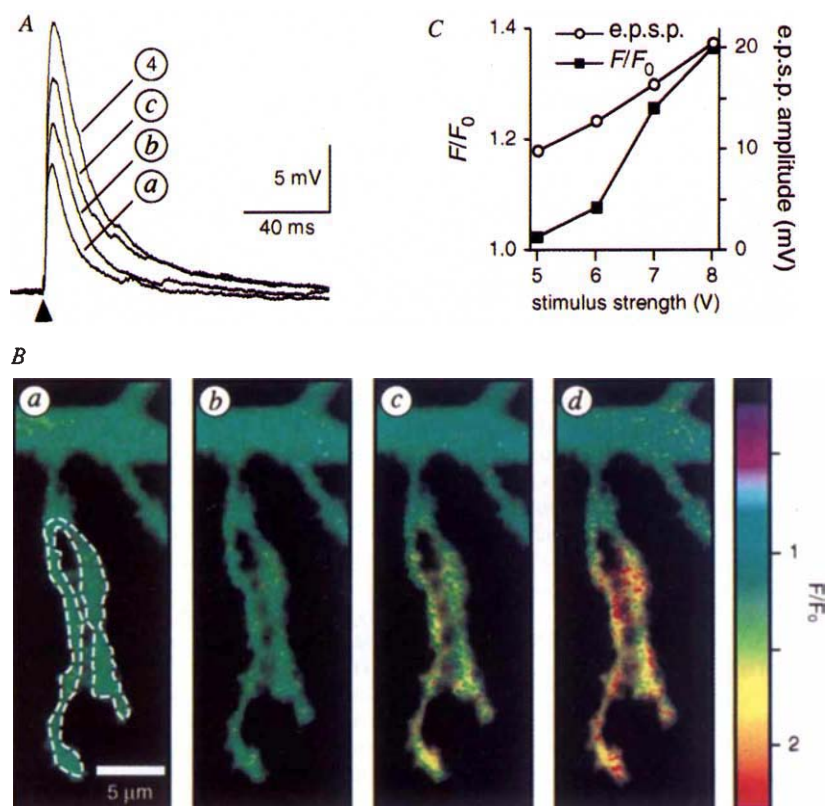


FIG. 4 Dependence of the subthreshold synaptic Ca^{2+} signals on e.p.s.p. amplitude. **A**, Superposed e.p.s.ps resulting from a stimulation pulse (100 μs) of increasing voltage. The increase in e.p.s.p. amplitude is due to an increase in number of activated afferent parallel fibres²². **B**, Pseudocolour images corresponding to the e.p.s.ps shown in **A**, as indicated. The dashed line within the left-most image indicates the region in which the fluorescence signals plotted in **C** were measured. **C**, Dependence of e.p.s.p. amplitudes and corresponding increases in fluorescence on stimulus strength.



Received 4 July; accepted 16 November 1994.

1. Eccles, J. C. *The Physiology of Synapses* 4-5, 101-121 (Springer, Berlin, 1964).
2. Bourne, H. R. & Nicoll, R. *Cell* **72**, 65-75 (1993).
3. Sugimori, M. & Llinás, R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5084-5088 (1990).
4. Midtgaard, J., Lasser-Ross, N. & Ross, W. N. *J. Neurophysiol.* **70**, 2455-2469 (1993).
5. Miyakawa, H., Lev-Ram, V., Lasser-Ross, N. & Ross, W. N. *J. Neurophysiol.* **68**, 1178-1189 (1992).
6. Llinás, R. R. & Hess, R. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2520-2523 (1976).
7. Llinás, R. R. & Sugimori, M. *J. Physiol.* **305**, 197-213 (1980).
8. Ross, W. N. & Werman, R. *J. Physiol.* **389**, 319-336 (1987).
9. Tank, D. W., Sugimori, M., Connor, J. A. & Llinás, R. R. *Science* **242**, 773-777 (1988).
10. Palay, S. L. & Chan-Palay, V. *Cerebellar Cortex* 79-85 (Springer, Berlin, 1974).
11. Müller, W. & Connor, J. A. *Nature* **354**, 73-76 (1991).
12. Neher, E. & Augustine, G. J. *J. Physiol.* **450**, 273-301 (1992).
13. Schneggenburger, R., Zhou, Z., Konnerth, A. & Neher, E. *Neuron* **11**, 133-143 (1993).
14. Murphy, T. H., Baraban, J. M., Wier, W. G. & Blatter, L. A. *Science* **263**, 529-532 (1994).
15. Alford, S., Frenguelli, B. G., Schofield, J. G. & Collingridge, G. L. *J. Physiol.* **469**, 693-716 (1993).
16. Bashir, Z. I. et al. *Nature* **363**, 347-350 (1993).
17. Hollmann, M., Hartley, M. & Heinemann, S. *Science* **252**, 851-853 (1991).
18. Hillman, D. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7076-7080 (1991).
19. Usowicz, M. M., Sugimori, M., Cherksey, B. & Llinás, R. R. *Neuron* **9**, 1185-1199 (1992).
20. Llano, I., Marty, A., Armstrong, C. M. & Konnerth, A. *J. Physiol.* **434**, 183-213 (1991).
21. Ito, M. *The Cerebellum and Neural Control* 36-37 (Raven, New York, 1984).
22. Konnerth, A., Llano, I. & Armstrong, C. M. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2662-2665 (1990).
23. Barbour, B. *Neuron* **11**, 759-769 (1993).
24. Ito, M. *Curr. Opin. Neurobiol.* **1**, 616-620 (1991).
25. Konnerth, A., Dreessen, J. & Augustine, G. J. *Proc. natn. Acad. Sci. U.S.A.* **89**, 7051-7055 (1992).
26. Kano, M., Rexhausen, U., Dreessen, J. & Konnerth, A. *Nature* **356**, 601-604 (1992).
27. Llinás, R. R. & Walton, K. D. in *The Synaptic Organization of the Brain* (ed. Shepherd, G. M.) 214-245 (Oxford Univ. Press, New York, 1990).
28. Edwards, F., Konnerth, A., Sakmann, B. & Takahashi, T. *Pflügers Arch.* **414**, 600-612 (1989).
29. Eilers, J., Schneggenburger, R. & Konnerth, A. in *Single Channel Recording* 2nd edn (eds Sakmann, B. & Neher, E.) 213-229 (Plenum, New York, 1995).

ACKNOWLEDGEMENTS. We thank G. Callewaert and J. Kauer for comments on the manuscript and N. Wilhelm for technical help. The work was supported by grants of the Bundesministerium für Forschung und Technologie, the Deutsche Forschungsgemeinschaft to A.K., and a Human Frontier Science Program grant to A.K. and G.J.A.

Neuregulins are concentrated at nerve-muscle synapses and activate ACh-receptor gene expression

Sangmee Ahn Jo*, Xuejun Zhu*†, Mark A. Marchionni† & Steven J. Burden*

* The Center for Blood Research, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02115, USA
 † Biology Department, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
 ‡ Cambridge NeuroScience Inc., One Kendall Square, Building 700, Cambridge, Massachusetts 02139, USA

Two different signalling pathways mediate the localization of acetylcholine receptors (AChRs) to synaptic sites in skeletal muscle. The signal for one pathway is agrin, a protein that triggers a redistribution of previously unlocalized cell surface AChRs to synaptic sites¹. The signal for the other pathway is not known, but this signal stimulates transcription of AChR genes in myofibre nuclei near the synaptic site². Neuregulins, identified originally as a potential ligand for erbB2 (Neu differentiation factor, NDF)³, stimulate proliferation of Schwann cells (glial growth factor, GGF)⁴, increase the rate of AChR synthesis in cultured muscle cells (AChR-inducing activity)⁵ and are expressed in motor neurons^{4,5}. These results raise the possibility that neuregulin is the signal that activates AChR genes in synaptic nuclei. Here we show that neuregulin activates AChR gene expression in C2 muscle cells and that the neuregulin response element in the AChR δ -subunit gene is contained in the same 181 base pairs that confer synapse-specific expression in transgenic mice. We use antibodies to show that neuregulins are concentrated at synaptic sites and that, like the extracellular signal that stimulates synapse-specific expression,

neuregulins remain at synaptic sites in the absence of nerve and muscle. We show that C2 muscle cells contain erbB2 and erbB3 messenger RNA but little or no erbB4 mRNA, and that neuregulin stimulates tyrosine phosphorylation of erbB2 and erbB3, indicating that neuregulin signalling in skeletal muscle may be mediated by a complex of erbB2 and erbB3.

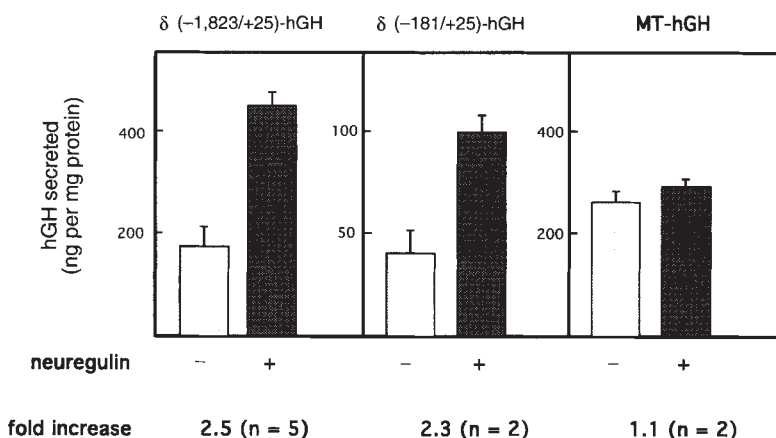
The AChR subunit genes, including the δ -subunit gene, are transcribed selectively in myofibre nuclei that are located near the synaptic site⁶⁻⁸. Studies in transgenic mice have shown that the *cis*-acting sequences for synapse-specific, electrical-activity-dependent and muscle-specific expression of the δ -subunit gene are contained in 181 base pairs (bp) of 5' flanking DNA⁹. To determine whether neuregulin activates transcription of the δ -subunit gene, we treated C2 muscle cells that were stably transfected with gene fusions between the 5' flanking region of the AChR δ -subunit gene and the human growth hormone (hGH) gene with recombinant neuregulin. Figure 1 shows that neuregulin causes a 2.5-fold increase in hGH secretion from C2 muscle cells stably transfected with an AChR δ (-1,823/+25)-hGH or an AChR δ (-181/+25)-hGH gene fusion. In contrast, neuregulin has no effect on the amount of hGH secreted from C2 muscle cells containing a metallothionein promoter-hGH gene fusion (Fig. 1). Thus, neuregulin stimulates transcription of the AChR δ -subunit gene, and the neuregulin response element is contained in the same 181 bp that confer synapse-specific expression in transgenic mice.

Selective expression of AChR genes in the synaptic nuclei of muscle is thought to be activated by a signal provided locally at synaptic sites^{2,6-8}. To determine whether neuregulin is present at synaptic sites, we prepared antibodies against neuregulin for use in immunocytochemical experiments. Multiple isoforms of neuregulin are generated by alternative splicing¹⁰, raising the possibility that different neuregulin isoforms may have different cellular and/or subcellular distributions. Therefore, we considered it important to raise antibodies that are not isoform-specific but would recognize most, if not all, isoforms of neuregulin. In this regard, we prepared antibodies against a peptide sequence within the epidermal growth factor (EGF) domain of neuregulin which is conserved among the different neuregulin isoforms but not among other proteins containing EGF domains^{3,11}. Figure 2 shows that the anti-peptide antibodies stain synaptic sites in frozen sections of skeletal muscle. Synaptic sites are also stained in unpermeabilized whole mount preparations of muscle (data not shown), indicating that the EGF domain of synaptic neuregulin is exposed extracellularly (see also Fig. 3g). Less intense antibody staining is observed in motor axons of intramuscular nerves (Fig. 2a). Antibody staining is not observed with pre-immune sera (Fig. 2c) and is inhibited by preincubation of immune sera with the peptide (Fig. 2e). Thus, neuregulin is concentrated at synaptic sites.

We do not know which neuregulin isoform is present at synaptic sites. Immunocytochemistry shows that at least one, and perhaps several, neuregulin isoforms are present at synaptic sites. As antibodies against an N-terminal peptide in glial growth factor (GGF2)⁴, which is unique to GGF2, stain intramuscular axons but not synapses (data not shown), not all neuregulin isoforms, however, are present at synaptic sites.

The signal that activates synapse-specific gene expression remains at synaptic sites following denervation^{6,12}, and previous studies have indicated that this signal is located in the synaptic basal lamina¹³⁻¹⁵. To determine whether neuregulin remains at synaptic sites following denervation, we stained denervated muscle with antibodies to neuregulin. Figure 3 shows that neuregulin remains at denervated synaptic sites. In contrast, components of the nerve terminal, including the synaptic vesicle protein SV2, are removed from synaptic sites following denervation (Fig. 3c, e). Figure 3 also shows that neuregulin remains at original synaptic sites following damage to nerve and muscle (Fig. 3g), indicating that neuregulin is associated with the synaptic basal lamina. The intensity of anti-neuregulin staining is reduced

FIG. 1 Neuregulin activates AChR δ -subunit gene expression. Neuregulin stimulates hGH expression from C2 muscle cells stably transfected with an AChR $\delta(-1,823/+25)$ -hGH or an AChR $\delta(-181/+25)$ -hGH gene fusion but not from C2 cells stably transfected with a metallothionein-hGH (MT-hGH) gene fusion. **METHODS.** C2 cells were cotransfected with pSV2-neo and an AChR δ subunit-hGH or a MT-hGH gene fusion, and pooled, stably transfected cells were isolated after selection in G418 (ref. 22). C2 myoblasts and myotubes were grown as described previously²², except that the growth medium was supplemented with chick embryo extract (0.5%). C2 myotubes were treated for 24 h with neuregulin (5 nM purified recombinant GGF2)^{4,23}, and the amount of hGH secreted into the medium was measured by a radioimmunoassay²². Mean $\pm$ s.e.m. secretion from two cultures treated with neuregulin and from four untreated cultures are shown. Mean increase in hGH expression from multiple experiments (n , number of repetitions for each experiment) is also indicated. Neuregulin increased the level of endogenous δ -subunit



mRNA 2.2-fold, similar to the effect of AChR-inducing activity in primary mouse myotubes²⁴.

after denervation, suggesting that neuregulin is also associated with nerve terminals. Because our antibodies recognize the EGF domain of neuregulin, which is sufficient for signalling¹¹, neuregulin that is present in the synaptic basal lamina is likely to be functional. The persistence of neuregulin at original synaptic sites in the absence of nerve and muscle is consistent with the idea that neuregulin is contained in the synaptic basal lamina and is the signal that activates synapse-specific gene expression.

One neuregulin isoform is secreted from cells⁴, whereas other neuregulin isoforms, which contain a single putative transmembrane domain, are thought to be membrane-associated^{3,5,10,11}. A soluble signalling domain might be released from membrane-associated neuregulin as a consequence of proteolytic processing^{3-5,10,11}. Thus, the basal lamina form(s) of neuregulin could be secreted from nerve terminals or derived by proteolytic processing of an isoform associated with the nerve terminal membrane.

Previous studies have indicated that erbB3 and erbB4 but not erbB2 bind neuregulin^{16,17}. Skeletal muscle is thought to express erbB4 (ref. 18) but not erbB3 mRNA¹⁹; potential expression of erbB2 in skeletal muscle has not been studied. We used an RNase protection assay to measure the abundance of erbB2, erbB3 and erbB4 mRNA in C2 muscle cells and in innervated and denerv-

ated skeletal muscle. Figure 4 shows that erbB2 and erbB3 are expressed in innervated and denervated adult skeletal muscle and in C2 myoblasts and myotubes. erbB3, unlike erbB2, is induced during muscle differentiation (Fig. 4a). Although low levels of erbB4 mRNA are detectable in innervated and denervated adult skeletal muscle, erbB4 expression is not detectable in C2 myoblasts and myotubes (Fig. 4a), suggesting that Schwann cells²⁰, rather than myofibres, are the source of erbB4 expression in skeletal muscle.

Because erbB3 is reported to lack kinase activity, it is thought that signalling through erbB3 requires association of erbB3 with another member of the EGF receptor family²¹. Neuregulin is not thought to bind erbB2 but to stimulate tyrosine phosphorylation of erbB2 when erbB2 is coexpressed with erbB3 or erbB4^{16,18}. Because erbB4 is not expressed in C2 myotubes and because the low level of erbB4 expression in skeletal muscle may be attributed to Schwann cells, we favour the idea that erbB3 is the predominant neuregulin receptor in muscle cells and that a complex of erbB2 and erbB3 mediate neuregulin signalling in skeletal muscle.

To determine whether neuregulin activates erbB2 and/or erbB3 in skeletal muscle cells, we treated C2 myotubes with neuregulin, immunoprecipitated erbB2 or erbB3, and deter-

FIG. 2 Neuregulin is concentrated at nerve-muscle synapses. Antibodies against neuregulin stain neuromuscular synapses (a, long arrows), which were identified by tetramethylrhodamine (TMR)-labelled α -bungarotoxin (b, d, f). Less intense antibody staining is also detected within axons of intramuscular nerves (a, short arrows); the myelinating Schwann cells are not stained. Antibody staining is not observed with preimmune serum (c) and is inhibited by pre-incubation of the immune sera with the 13-amino-acid peptide (15 μ M) (e). Scale bar, 15 μ m. **METHODS.** A 13-amino-acid peptide (amino acids 197-209 of Neu differentiation factor)³ was coupled to a MAP carrier core (Research Genetics), and rabbits were immunized with the conjugate. Frozen sections (10 μ m) of the rat sternomastoid muscle were stained with anti-neuregulin sera (diluted 1/1,500) followed by fluorescein-conjugated secondary antibodies (a, c, e) and TMR-labelled α -bungarotoxin (b, d, f). Similar staining was observed with antisera from two rabbits and in the soleus muscle.

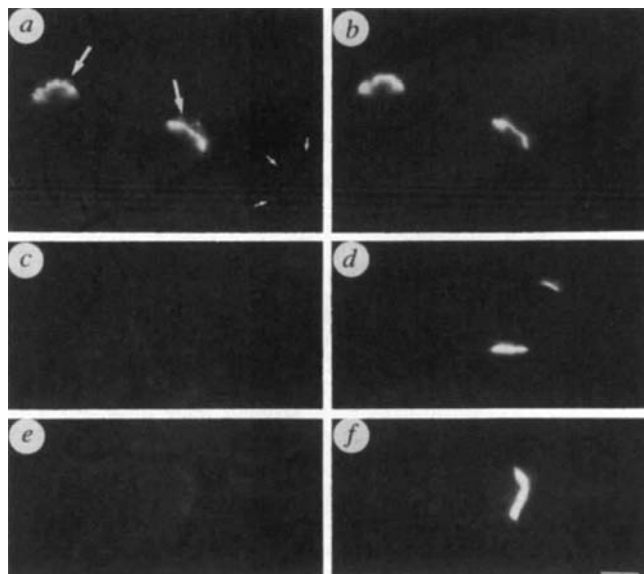


FIG. 3 Neuregulin remains at synaptic sites in the absence of nerve and muscle. **a**, Antibodies against neuregulin stain denervated neuromuscular synapses (arrows), which were identified by TMR- α -bungarotoxin (**b**, **d**, **f**). In contrast, antibodies against the synaptic vesicle protein, SV2 (ref. 25), stain synaptic sites in innervated (**c**, arrow) but not in denervated (**e**) muscle. Similarly, a monoclonal antibody (SMI31, Sternberger Monoclonals Inc.) to neurofilaments stains synaptic sites in innervated but not in denervated muscle (data not shown). Antibodies against neuregulin stain original synaptic sites (**g**), identified by s-laminin staining (**h**), in the absence of nerve and muscle. A low level of synaptic staining, similar to that illustrated in (**e**), was observed in the absence of fluorescein-conjugated secondary antibodies and was due to TMR fluorescence detected in the fluorescein-selective channel. Scale bar, 18 μ m.

METHODS. The nerve to the sternomastoid muscle was cut at the edge of the muscle, and frozen sections from 4-day denervated muscle were stained with antibodies and TMR- α -bungarotoxin (**b**, **d**, **f**) as in Fig. 2. Damaged muscle was stained with antibodies to neuregulin (**g**) and s-laminin (**h**) two days after damaging nerve and muscle as described previously¹³. s-laminin, a component of the synaptic basal lamina²⁶, was detected with rhodamine-conjugated secondary antibodies.

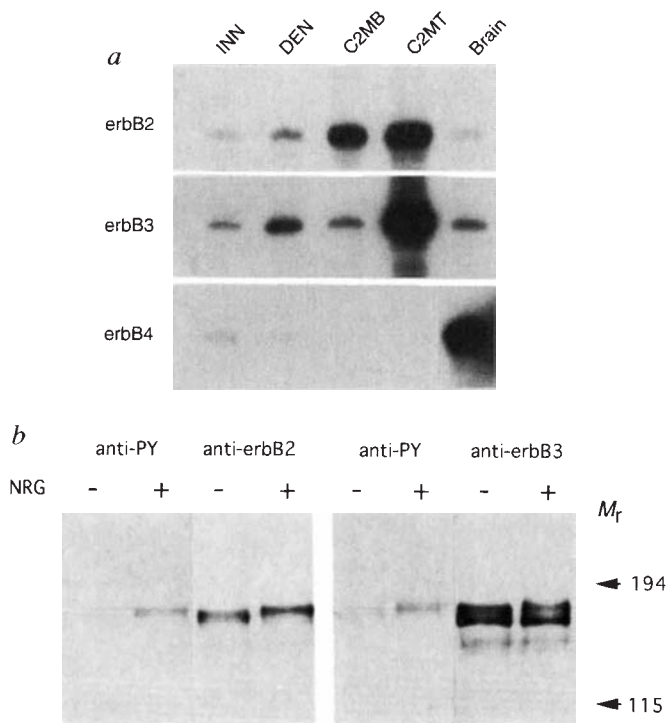
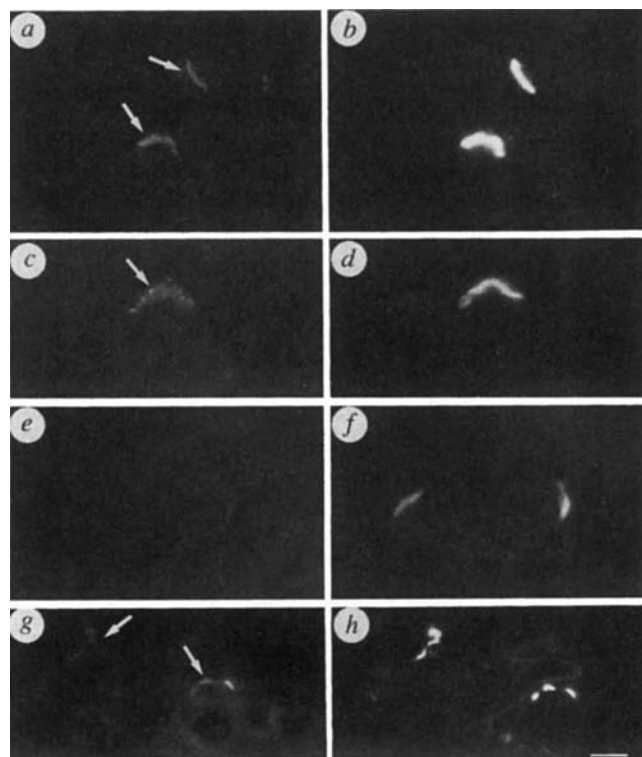


FIG. 4 Neuregulin stimulates tyrosine phosphorylation of erbB2 and erbB3 in skeletal muscle cells. **a**, erbB2 and erbB3 mRNA is expressed in C2 muscle cells. erbB2 mRNA is expressed in innervated (INN) and denervated (DEN) adult mouse skeletal muscle and in C2 myoblasts (C2MB) and C2 myotubes (C2MT), and erbB3 mRNA is expressed in innervated and denervated mouse skeletal muscle (expression of erbB2 and erbB3 increases two- to threefold following denervation; two experiments each). Expression of erbB3 is activated during muscle differentiation, as erbB3 mRNA levels are low in C2 myoblasts and increase ~25-fold following myoblast fusion. The level of erbB3 mRNA in skeletal muscle is comparable to that in brain. erbB4 mRNA is expressed at low levels in innervated and denervated adult mouse skeletal muscle (expression decreases 20% following denervation; two experiments). Expression of erbB4 mRNA is not detected in C2 myoblasts or myotubes.

The level of erbB4 mRNA in skeletal muscle is low compared with that in brain. The level of erbB4 mRNA is ~10-fold less than erbB3 mRNA in adult muscle. **b**, Neuregulin (NRG) stimulates tyrosine phosphorylation of erbB2 and erbB3 in C2 myotubes. erbB2 or erbB3 were immunoprecipitated with antibodies specific for each receptor from untreated C2 myotubes (-) or C2 myotubes treated with neuregulin (+). Phosphotyrosine incorporation was visualized with antibodies to phosphotyrosine (anti-PY), and erbB2 and erbB3 expression was visualized with antibodies (anti-erbB2 or anti-erbB3) specific for each receptor. erbB2 and erbB3 migrate similarly in SDS-PAGE and have an apparent M_r of 180K.

METHODS. RNase protection assays were as described previously⁶. The rat erbB2 probe extends between nucleotides 3,245 to 3,529 (ref. 27). Because we used a rat erbB2 probe to measure the abundance of mouse erbB2 mRNA, we cannot make a direct comparison of the abundances of erbB2 mRNA with erbB3 and erbB4 mRNA in mouse muscle. We isolated a complementary DNA fragment encoding mouse erbB3; the mouse sequence corresponds to the region of human erbB3 cDNA extending between nucleotides 3,197 and 3,609 (ref. 28) and was used to synthesize an erbB3 RNA probe. The mouse erbB4 probe extends between nucleotides 3,360 and 3,573 (ref. 18). Gels were exposed to X-ray film overnight (erbB2, erbB3) or for three days (erbB4) at -80 °C with an intensifying screen. cDNA encoding amino acids 1 to 241 of heregulin $\beta 1^{11}$ was cloned into a baculovirus expression vector (PharMingen). SF9 cells were infected with the recombinant virus, and conditioned medium was collected three days following infection. C2 myotubes were treated for 10 min either with conditioned medium (diluted 1:1 with C2 fusion medium) from SF9 cells infected with wild-type virus (-) or from SF9 cells infected with the recombinant virus (+). Cells were lysed in RIPA buffer (20 mM HEPES, pH 7.4, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100, 1 mM EDTA, 1 mM orthovanadate, 2 mM leupeptin, 1 mM pepstatin, 1 mM pefabloc, 1 mM aprotinin) and erbB2 was immunoprecipitated with an monoclonal antibody (c-neu Ab3) against erbB2 (Oncogene Sciences); erbB3 was immunoprecipitated with rabbit antibodies (C17) against erbB3 (Santa Cruz Biotechnology). The immunoprecipitated proteins were collected on Protein A-Sepharose beads, fractionated by SDS-PAGE and transferred to nitrocellulose. Western blots were probed either with an antibody (RC20H) against phosphotyrosine (Transduction Laboratories), antibodies (6-2987) against erbB2 or antibodies (C17) against erbB3. The specificity of the anti-erbB antibodies was determined by showing that anti-erbB2 antibodies, but not anti-erbB3 antibodies, labelled a 180K protein immunoprecipitated with anti-erbB2. Similarly, anti-erbB3 antibodies, but not anti-erbB2 antibodies, labelled a doublet of 180K immunoprecipitated with anti-erbB3. Antibody binding was detected by enhanced chemiluminescence.

mined whether neuregulin induced tyrosine phosphorylation of erbB2 and erbB3. Figure 4b shows that neuregulin stimulates tyrosine phosphorylation of both erbB2 and erbB3 in C2 myotubes. Thus, erbB2 and erbB3 are likely to have a role in neuregulin-mediated signalling in skeletal muscle. □

Received 19 September; accepted 29 November 1994.

- McMahan, U. J. *Cold Spring Harb. Symp. quant. Biol.* **55**, 407–418 (1990).
- Burden, S. J. *Trends Genet.* **9**, 12–16 (1993).
- Wen, D. et al. *Cell* **69**, 559–572 (1992).
- Marchionni, M. A. et al. *Nature* **362**, 312–318 (1993).
- Falls, D. L. et al. *Cell* **72**, 801–815 (1993).
- Simon, A. M., Hoppe, P. & Burden, S. J. *Development* **114**, 545–553 (1992).
- Klarsfeld, A. et al. *EMBO J.* **10**, 625–632 (1991).
- Sanes, J. R. et al. *Development* **113**, 1181–1191 (1991).
- Tang, J., Jo, S. A. & Burden, S. J. *Development* **120**, 1799–1804 (1994).
- Mudge, A. W. *Curr. Biol.* **3**, 361–364 (1993).
- Holmes, W. E. et al. *Science* **256**, 1205–1210 (1992).
- Gundersen, K., Sanes, J. R. & Merlie, J. P. *J. Cell Biol.* **123**, 1535–1544 (1993).
- Jo, S. A. & Burden, S. J. *Development* **115**, 673–680 (1992).
- Brenner, H. R., Herczeg, A. & Slater, C. R. *Development* **116**, 41–53 (1992).

- Goldman, D., Carlson, B. M. & Staple, J. *Neuron* **7**, 649–658 (1991).
- Carraway, K. L. et al. *J. Biol. Chem.* **269**, 14303–14306 (1994).
- Plowman, G. D. et al. *Nature* **366**, 473–475 (1993).
- Plowman, G. D. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1746–1750 (1993).
- Kraus, M. H., Issing, W., Miki, T., Popescu, N. C. & Aaronson, S. A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9193–9197 (1989).
- Lai, C. & Lemke, G. *Neuron* **6**, 691–704 (1991).
- Carraway, K. L. & Cantley, L. C. *Cell* **78**, 5–8 (1994).
- Simon, A. M. & Burden, S. J. *Mol. cell. Biol.* **13**, 5133–5140 (1993).
- Shah, N. M., Marchionni, M. A., Isaacs, I., Stroobant, P. & Anderson, D. J. *Cell* **77**, 349–360 (1994).
- Martinou, J.-C., Falls, D. L., Fischbach, G. D. & Merlie, J. P. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7669–7673 (1991).
- Buckley, K. M. & Kelly, R. B. *J. Cell Biol.* **100**, 1284–1294 (1985).
- Sanes, J. R. & Chiu, A. Y. *Cold Spring Harb. Symp. quant. Biol.* **48**, 667–678 (1983).
- Bargmann, C. I., Hung, M. C. & Weinberg, R. A. *Nature* **319**, 226–230 (1986).
- Plowman, G. D. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4905–4909 (1990).

ACKNOWLEDGEMENTS. The first two authors contributed equally to this work. We thank C. Lai for mouse erbB4 cDNA, K. Buckley for monoclonal antibody to SV2, J. Sanes for monoclonal antibody (D5) to s-laminin, A. Simon for C2 muscle cells stably transfected with an AChR $\delta(-1,823/+25)$ -hGH gene fusion, I. Isaacs for purified recombinant neuregulin, R. Schatzman for antibodies (6-2987) to erbB2, K. Carraway for help with the baculovirus expression system; and J. Yeadon and R. Lehmann for comments on the manuscript. This work was supported by the NIH, the Muscular Dystrophy Association and Cambridge Neuroscience Inc.

Minus-end-directed motion of kinesin-coated microspheres driven by microtubule depolymerization

Vivian A. Lombillo, Russell J. Stewart* & J. Richard McIntosh†

Department of Molecular, Cellular and Developmental Biology,
University of Colorado, Boulder, Colorado 80309, USA
* Rowland Institute for Science, Cambridge,
Massachusetts 02142, USA

DYNAMIC changes in microtubule (MT) length have long been thought to contribute to intracellular motility¹. Both the polymerization² and depolymerization^{3–5} of tubulin have been shown to do work *in vitro*, but the biochemical complexity of objects moved, such as chromosomes, has complicated the identification of proteins that couple MT dynamics with motility. Work with MTs grown from and tethered to pellicles of lysed *Tetrahymena* has shown that disassembly-dependent movement of chromosomes *in vitro* can be inhibited with antibodies against the motor domain of kinesin⁶. To study proteins that can function in disassembly-dependent motion, we have refined this motility assay, replacing chromosomes with protein-coated latex microspheres. We report here the ability of several enzymes, including kinesin, to support *in vitro* motility of latex microspheres on disassembling MTs (Fig. 1a). The polarity of kinesin's motor activity can be reversed by MT disassembly and interactions between a motor and a MT end can either slow or speed the rate of tubulin depolymerization.

Microspheres coated with HeLa cell kinesin move with the end of disassembling MTs. This kinesin supported robust, ATP-dependent gliding of MTs when adsorbed to glass, demonstrating that its motor activity was preserved (not shown). When bound to pellicle-initiated MTs in the presence of Mg-ATP, kinesin-coated microspheres moved toward the MT plus-ends (Fig. 1b). The average rate for this 'conventional', ATP-dependent motility was $53.4 \pm 11.8 \mu\text{m min}^{-1}$ ($n=11$), a value typical for kinesins⁷.

The direction of motor motion can be reversed when these kinesin-associated microspheres are bound to MTs that are depolymerizing. In a buffer similar to that used in ref. 8 to study diffusional motions of MTs bound to glass by flagellar dynein, MTs begin to depolymerize within minutes after dilution, and

kinesin-coated spheres follow shortening MTs in toward their minus-ends at the pellicle surface (Fig. 1c, d; Table 1). The average rate of depolymerization-dependent motility of these spheres was $22.9 \pm 6.4 \mu\text{m min}^{-1}$ ($n=26$). When 1 mM Mg-ATP was added to the dilution buffer, microspheres still moved towards the MT minus-ends, but their speed increased by a factor of about 2 (Table 1). This increase probably derives from the decreased affinity between MTs and kinesin that is caused by ATP⁷.

Because the dilution buffer that promotes motility in our assay also allows MTs tethered by motor enzymes to show one-dimensional diffusion⁸, disassembly-dependent motility may require not only the depolymerization of tubulin subunits but also a suitably weak association between the coupling protein and the MT surface. To test this idea we looked for motion of microspheres coated with two proteins known to bind weakly to MTs and to support one-dimensional diffusion over their surfaces: flagellar dynein⁸ and a genetically engineered kinesin-like protein⁹. In the presence of ATP-VO₃³⁻, flagellar dynein-coated microspheres moved towards the minus-end of disassembling MTs. Their velocity was similar to that measured for analogous motions of kinesin-coated spheres along disassembling MTs in the presence of ATP (Table 1).

The second weakly binding molecule tested was NK350, a newly engineered chimera that consists of the motor domain of *Drosophila* kinesin heavy chain and the stalk domain of the *Drosophila* kinesin-like protein, NCD⁹ (Fig. 2a). In a standard MT gliding assay, NK350 adsorbed to glass will bind MTs, but the polymers do not show conventional movement; they simply shuttle back and forth in one dimension (Fig. 2b, c). The mean square displacement of MTs tethered by NK350 increased linearly with time, as predicted for a diffusion-based process (Fig. 2d). NK350 supported minus-end-directed depolymerization-dependent motion of microspheres in tubulin assembly buffer (see Fig. 1 legend and Table 1). Spheres moved at constant velocities interrupted by occasional pauses (Fig. 2e), a pattern reminiscent of the chromosome and vesicle motility described previously^{4,5}. However, NK350-coated spheres, together with the visible ends of disassembling MTs, moved significantly faster ($v=186 \pm 84 \mu\text{m min}^{-1}$, $n=45$) than when the spheres were coated with the other motors studied.

The normal speed of MT disassembly following rapid subunit dilution in PME is $\sim 23 \mu\text{m min}^{-1}$ (Table 1 and ref. 10). The speed of disassembly-dependent movement of NK350-coated spheres is therefore surprisingly fast, fast enough to be inconsistent with models which postulate that microsphere motility results from a passive, one-dimensional diffusion of the coupling molecules that is simply biased in direction by MT shortening. Results with NK350 suggest that a 'coupling' motor can aid subunit loss from the MT plus-end. Soluble NK350 does not,

† To whom correspondence should be addressed.

FIG. 1 *a*, *In vitro* assay for depolymerization-drive motion. *Tetrahymena* pellicles were bound to a coverslip and used to initiate MT growth in 25 μ M tubulin. A perfusion chamber was constructed using this coverslip and a microscope slide⁵. Latex microspheres were coated with the protein to be tested, then added to the chamber where they bound to MTs. MTs were depolymerized by subunit dilution or CaCl_2 and sphere motility was recorded. *b*, Microsphere motility powered by ATP-dependent motor activity. Kinesin-coated latex microspheres bound MTs in PME buffer (100 mM PIPES, pH 6.9, 2 mM MgSO_4 , 1 mM EGTA, 0.5 mM DTT), containing 1 mM Mg-ATP and 25 μ M tubulin. These spheres moved along toward the MT plus-ends. *c*, ATP-independent, depolymerization-driven movement. The concentration of tubulin was reduced to <10 μ M with 10 mM Tris, pH 7.5, 40 mM KCl, 2 mM CaCl_2 , 1 mM DTT (no ATP). Kinesin-coated spheres moved with the ends of MTs towards their minus ends at the pellicle surface. *d*, Kinesin-coated microspheres moving with the ends of MTs grown from an axoneme as they shortened under the same buffer conditions used in *c*. The single sphere moves proximally first, followed by two others attached to a different MT bundle (white arrow). In *b*–*d*, arrowheads indicate MT bundles associated with the spheres. Black arrows indicate direction of flow. Times are given in seconds. Sphere diameter, 1 μ m.

METHODS. Assay components and procedures used were previously described in detail⁵. Latex microspheres (1 μ m; Polysciences Inc.) were coated with purified proteins and tested for motility. Kinesin was prepared from HeLa cells, using a modification of the protocol described in ref. 14, then mixed with enough uncharged latex microspheres to yield ~50 motors per sphere and left on ice for 30–120 min. Spheres were then pelleted (16,000g for 1 min) and rinsed twice in PME. Coated spheres were allowed to bind to pellicle-initiated MTs for 5 min. In *b*, the chamber contained 25 μ M tubulin and 1 mM Mg-ATP in PME; motility was observed within minutes. The chamber in *c* was initially identical to that in *b*, but no ATP was present and tubulin was diluted with the Tris buffer described above. MTs are difficult to discern in *b* and *c* because the focal plane is >10 μ m from the coverslip. The axonemal MT bundle in *d* is much clearer, simply because it is closer to the coverslip. $T=32\pm2^\circ\text{C}$ throughout all experiments.

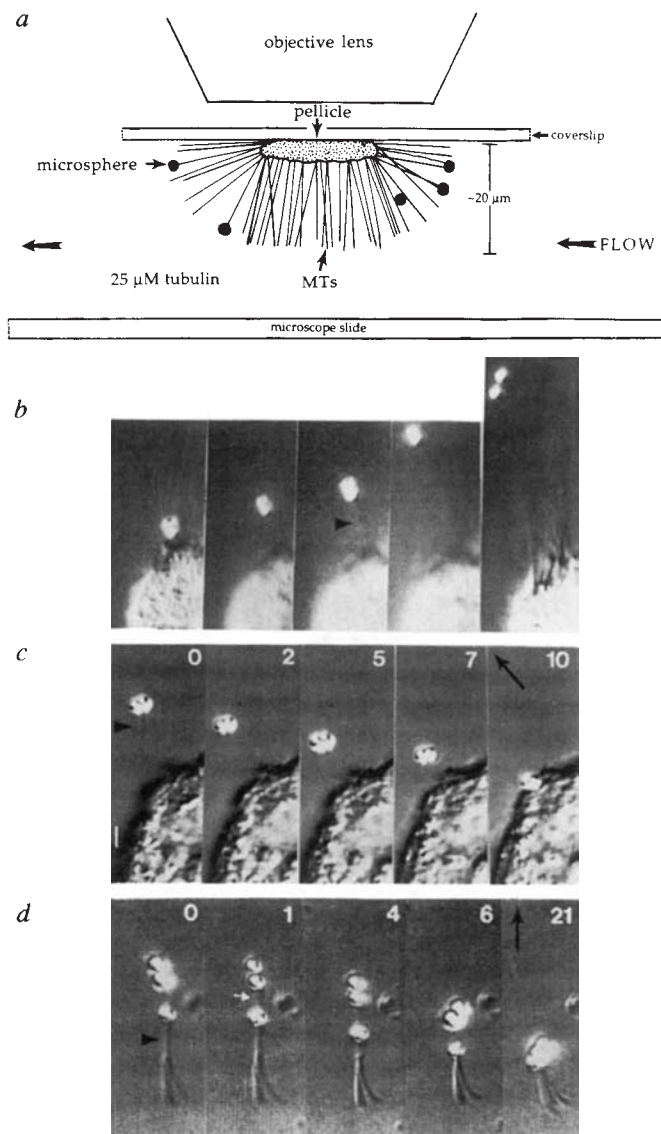


TABLE 1 Motility of protein-coated microspheres

Buffer	MT depolymerization rate ($\mu\text{m min}^{-1}$)	Bound protein	Assay condition	Polarity	Microsphere velocity ($\mu\text{m min}^{-1}$)
1. PME	23 ($n=4$)	Kinesin (HeLa)	+ATP, high tubulin	+	53.4 (s.d. 11.8, $n=11$)
2. Tris + CaCl_2		Kinesin (HeLa)	+ATP, low tubulin	–	58.1 (s.d. 14.7, $n=15$)
3. Tris + CaCl_2	148 ($n=3$)	Kinesin (HeLa)	–ATP, low tubulin	–	22.9 (s.d. 6.4, $n=26$)
4. Tris + CaCl_2		Flagellar dynein (sea urchin)	+ATP & VO_4^{3-} , low tubulin	–	56.1 (s.d. 17.9, $n=10$)
5. PME		NK350 chimaera	–ATP, high tubulin	n/a	0.0 ($n=45$)
6. PME	23 ($n=4$)	NK350 chimaera	+/- ATP, low tubulin	–	186.0 (s.d. 84.0, $n=45$)
7. PME + CaCl_2	84 ($n=5$)				
8. Tris	31 ($n=5$)				

The rates of microsphere movement and/or MT depolymerization depend on the buffer used for tubulin dilution. Exact compositions of Tris and PME are defined in the text. CaCl_2 , when present, 2 mM. Assay condition refers to the buffer introduced into the perfusion chamber at the beginning of the experiment: $\pm$ ATP (1.0 mM or 0.0 mM CaCl_2 , when present, 2 mM added Mg-ATP); high/low tubulin (25 μ M or during tubulin dilution to 0; VO_4^{3-} , 100 μ M). Polarity implies movement toward the plus (+) or minus (–) end of the MT. Velocities are given as means and standard deviations with the number of observations. Lines 7 and 8 describe mean MT depolymerization rates under conditions relevant to interpretation of the results in lines 1–6. Axonemes prepared from *Chlamydomonas*⁵ were adsorbed to glass coverslips and assembled into a chamber as for Fig. 1. MTs were grown from axonemes precisely as from pellicles, then the tubulin concentration was dropped by perfusing up to 5 chamber volumes of the different buffers tested. MTs were imaged with differential interference contrast optics and their behaviour was recorded. Positions of MT plus-ends were recorded during perfusion, directly from the video monitor and mean rates of disassembly were calculated from these length changes. Sea urchin sperm flagellar dynein was prepared as described in ref. 13.

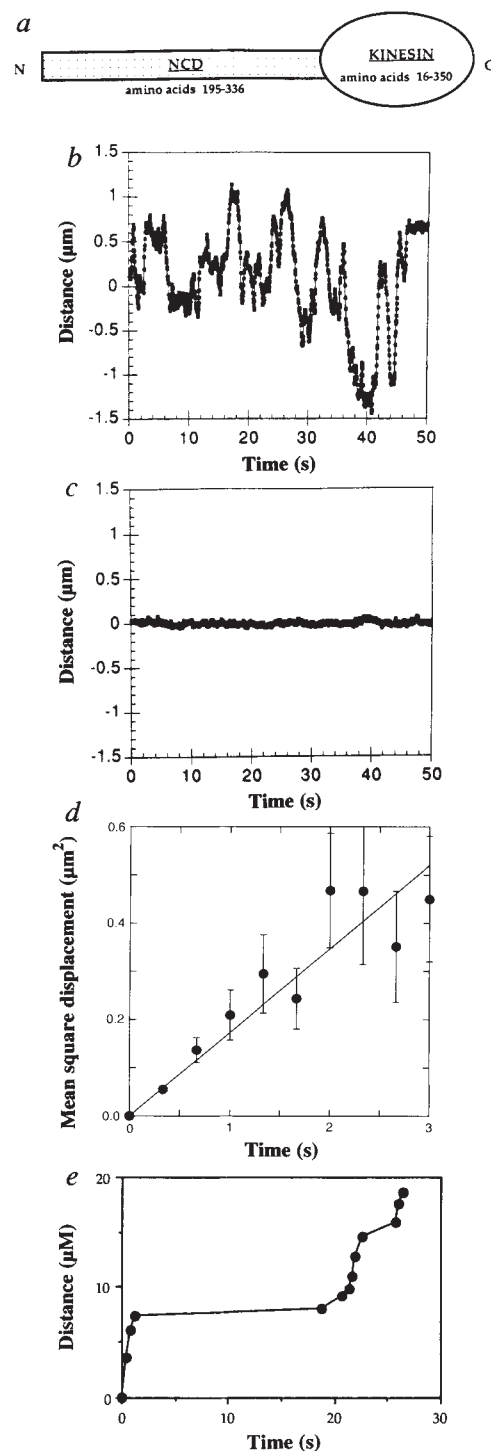


FIG. 2 Structure and behaviour of the recombinant chimaera, NK350. *a*, NK350 comprises the stalk domain of NCD coupled to the motor domain of the *Drosophila* kinesin heavy chain. MTs bound to glass by NK350 display 1-D diffusion. Displacements parallel to (*b*) and perpendicular to (*c*) the MT axis are plotted over time. *d*, Mean squared displacement along the MT axis is plotted against time. Correlation coefficient for straight line = 0.985. The observed diffusion rate ($D = 1.7 \pm 0.1 \times 10^{-9} \text{ cm}^2 \text{ s}^{-1}$) is slightly faster than for MTs bound by dynein⁸, implying a weaker binding. ATP neither affects this behaviour nor induces directed movement. NK350 does hydrolyse ATP, but the ATPase activity shows almost no MT activation (not shown). *e*, NK350 supports depolymerization-dependent motion of microspheres in the absence of ATP. Sphere position is plotted against time after tubulin dilution. The pauses in motility seen here are characteristic both of MT depolymerization¹⁰ and of chromosome motion *in vitro* driven by MT depolymerization⁴.

METHODS. The plasmid encoding NK350 was constructed by PCR amplification of bases 689–1,133 of *Drosophila* NCD and bases 370–576 of *Drosophila* kinesin heavy chain (khc). An *Xma*I restriction site was added to the NCD coding primer sequence. The noncoding NCD primer sequence was modified to change the NCD amino-acid sequence FCRIR (single-letter code) to the khc amino-acid sequence, VCRFR. The resulting DNA fragments were mixed and re-amplified with primers from the sequence ends. The resulting chimaeric DNA fragment was digested with *Xma*I and *Acc*I to create pGEX-NK350. The NK350 glutathione-S-transferase (GST) fusion protein was partially purified and assayed for MT motility *in vitro* as described previously⁹. MTs were bound but not moved by the intact GST-NK350 fusion protein adsorbed to a coverslip; however, NK350 released from GST by thrombin cleavage allowed MTs to undergo 1-D diffusion. Spheres were coated with NK350 protein as described for the kinesin experiments above. All assays with NK350 were in PME buffer without ATP or CaCl_2 .

however, accelerate tubulin depolymerization *in vitro* (not shown). Apparently, the chimaera must be structure-bound to speed subunit loss.

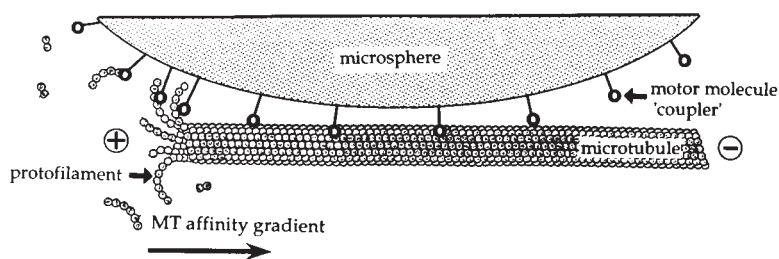
We also investigated whether cytoplasmic dynein prepared from HeLa¹¹ or *Dictyostelium*¹² can support motility of microspheres. Carboxylated spheres coated with cytoplasmic dynein initially stabilized MTs to which they were bound, but when the MTs eventually depolymerized, the spheres fell off the MTs and washed away in the flow, regardless of the dilution buffer used.

These results demonstrate that several but not all motor molecules can couple objects to the ends of shortening MTs. At least two of the effective molecules can diffuse along the MT lattice, so their movement with the end of a depolymerizing MT is noteworthy but not entirely unexpected. It is more surprising that conventional kinesin is also able to support depolymeriza-

tion-dependent motion of microspheres, particularly because the direction of movement is opposite to that induced by Mg-ATP with this enzyme on stable MTs. The ability of MT dynamics to change a motor's direction of motion is intriguing.

Rates of depolymerization-driven, ATP-independent motion of microspheres coated with both NK350 and kinesin suggest that molecules bound at MT ends can affect the rate of tubulin depolymerization. Because our preparations of NK350 and kinesin have no apparent effect on MT dynamics in solution, we infer that these motors must be associated with an object (for example, a microsphere) to affect depolymerization. It is easy to imagine how proteins, such as kinesin, may cap or stabilize the plus-end of a MT and consequently decrease the rate of subunit loss. The more surprising result of facilitated depolymerization with NK350-coated microspheres leads us to propose a model

FIG. 3 A model for the mechanism for depolymerization-driven motility of microspheres. Sphere-bound motor enzymes are proposed to interact with tubulin dimers positioned near the MT end. A gradient of MT affinity is suggested: motors bind most tightly to tubulin that is assembled into the straight protofilaments of the MT wall, at intermediate levels to tubulin near the ends of a shortening MT, and at lowest levels to soluble tubulin. This gradient moves the 'coupling' motor enzymes along the disassembling polymer towards the MT minus end. Movement is suggested to result from a thermally driven vibration that would allow either a linear or rotational sampling of nearby positions along MT protofilaments. This sampling could be achieved by a stepping of the motor enzyme, analogous to its conventional mechanochemistry, or by a dissociation-reassociation, influenced by buffer conditions. Protofilaments at the end of shortening MTs are known to splay out from the tubule axis¹⁵. Our model is drawn to reflect this property, which may be mechanistically significant for movement. The inter-



mediate affinity of the motor domain for curved protofilaments could lead to enhanced protofilament curving, destabilizing more tubulin subunits near the polymer end and thereby promoting MT depolymerization. Alternatively, coupling motors on the end of MTs could affect the rates of MT dynamics, for example by altering the GTPase activity of tubulin subunits¹⁶.

for the relationship between MT shortening and sphere motility (Fig. 3). *In vivo*, structure-bound motor enzymes may serve to couple MT dynamics to organelle motility. Such an activity is likely to be particularly important for MT function during mitosis, when MTs shorten and lengthen in association with chromosome movement. □

Received 15 June; accepted 15 November 1994.

1. Inoue, S., *J. Cell Biol.* **91**, 131s-147s (1981).
2. Miyamoto, H. & Hotani, H. *Proc. Taniguchi Int. Symp.* **14**, 220-242 (1988).
3. Koshland, D. E., Mitchison, T. J. & Kirschner, M. W. *Nature* **331**, 499-504 (1988).
4. Coue, M., Lombillo, V. A. & McIntosh, J. R. *J. Cell Biol.* **112**, 1165-1175 (1991).
5. Lombillo, V. A., Coue, M. & McIntosh, J. R. *Meth. Cell Biol.* **39**, 148-165 (1993).

6. Lombillo, V. A., Nislow, C., Yen, T. J., Gelfand, V. I. & McIntosh, J. R. *J. Cell Biol.* (in the press).
7. Vale, R. D. *et al. Cell* **42**, 39-50 (1985).
8. Vale, R. D., Soll, D. R. & Gibbons, I. R. *Cell* **59**, 915-925 (1989).
9. Stewart, R. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **90**, 5209-5212 (1993).
10. Walker, R. A. *et al. J. Cell Biol.* **107**, 1437-1448 (1988).
11. Vaisberg, G., Koonce, M. P. & McIntosh, J. R. *J. Cell Biol.* **123**, 849-858 (1993).
12. Koonce, M. P. & McIntosh, J. R. *Cell Motil. Cytoskel.* **15**, 51-62 (1990).
13. Porter, M. E. *et al. J. biol. Chem.* **263**, 6759-6771 (1988).
14. Hollenbeck, P. J. *J. Cell Biol.* **108**, 2335-2342 (1989).
15. Mandelkow, E. M., Milligan, R. & Mandelkow, E. J. *Cell Biol.* **114**, 977-991 (1991).
16. Davis, A., Carleton, R. S., Dougherty, C. A. & Farrell, K. W. *Science* **264**, 839-842 (1994).

ACKNOWLEDGEMENTS. We thank L. S. B. Goldstein for his support during the construction and early characterization of NK350 (done by R.J.S.), C. Nislow, E. M. Stone and L. Pillus for reading the manuscript, P. Furcinetti for computing assistance and C. Nislow, A. Hunt, E. Vaisberg and C. Chessick for discussion. This work was funded by a predoctoral grant from the NIGMS (to V.A.L.) and grants from the NIH (to R.J.S. and J.R.M.).

Structural polymorphism of bacterial adhesion pili

Esther Bullitt* & Lee Makowski†

* Department of Biophysics, Boston University School of Medicine, Boston, Massachusetts 02118-2394, USA

† Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306-3015, USA

BACTERIAL adhesion pili are designed to bind specifically and maintain attachment of bacteria to target cells. Uropathogenic P-pili are sufficiently mechanically resilient to resist the cleansing action of urine flow that removes most other bacteria¹. P-pili are 68 Å in diameter and ~1 µm long², and are composed of ~1,000 copies of the principal structural protein, PapA³. They are attached to the outer membrane by a minor structural protein, PapH⁴ and are terminated by an ~20 Å diameter fibrillus composed of PapK, PapE and PapF, which presents the host-binding adhesin PapG⁵⁻⁷. The amino-acid sequences of PapA^{3,8}, PapE⁹, and PapF⁹ are similar, with highly conserved C-termini being responsible for binding to PapD¹⁰⁻¹², the periplasmic chaperone. Our three-dimensional reconstruction indicates that pili are formed by the tight winding of a much thinner structure. A structural transition allows the pilus to unravel without depolymerizing, producing a thin, extended structure five times the length of the original pilus.

Electron micrographs of negatively stained P-pili from *Escherichia coli* (Fig. 1a) indicate that the pili are relatively rigid, rod-like structures. Figure 1b contains a power spectrum (square of the Fourier transform) of a micrograph of a negatively stained pilus. The sharp layer lines visible indicate that the pilus has a well ordered helical structure, confirmed by X-ray diffraction from oriented fibers of the pili². From the distribution of reflections on layer lines in the power spectra of 21 pili, we have determined the symmetry of P-pili to be 23 subunits in 7 turns

of the helix, or 3.28 subunits per turn (3.279 ± 0.003). The pitch of the helix is 24.9 Å, with a rise per subunit of 7.58 Å (compare with 3.3 subunits per turn, pitch 24.45 Å, and 7.42 Å rise per subunit from X-ray data²). This helical symmetry is essentially identical to that of type 1 pili from *E. coli*¹³. The hand of the helix was determined by depositing a platinum coating on the top surface of the pili by rotary shadowing. The three-start helix, pitch ~90 Å, is clearly left-handed in these micrographs (data not shown). From the helical symmetry of the pili, this observation defines the genetic, one-start helix as right-handed.

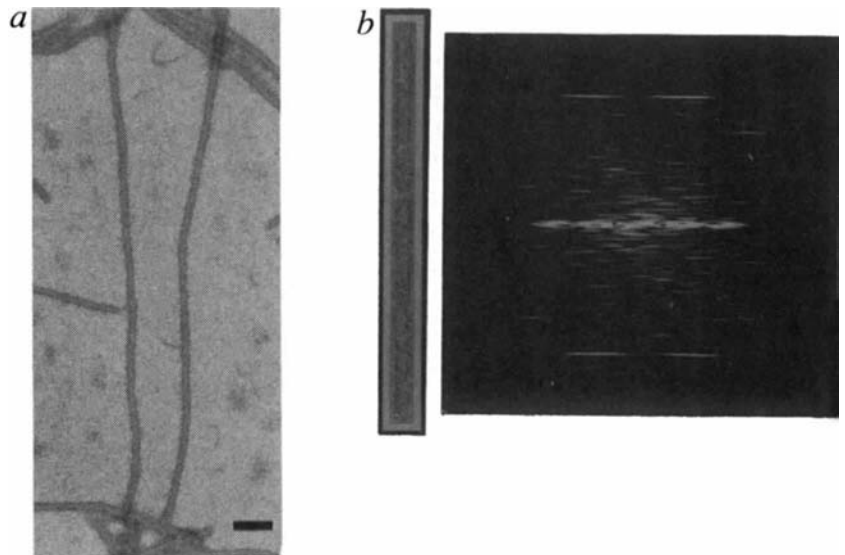
Figure 2 is a computer rendering of a three-dimensional reconstruction of the P-pili structure. Criteria for choosing filaments included in the reconstruction were straightness of the filaments and sharpness and symmetry of layer lines in the computed Fourier transforms. The contour level for the rendering in Fig. 2 was chosen so that the volume of each pilus subunit corresponds to that of a single PapA molecule. In the three-dimensional reconstruction, the pili have a maximum diameter of 68 Å, with a helical cavity of ~25 Å × 15 Å winding round the centre of the pilus. The centre of the cavity is displaced 5 Å from the helix axis, follows the one-start helix of the filament, and is connected to the outside of the pilus by a set of radial channels.

Boundaries between pili subunits are not clearly defined by the three-dimensional reconstruction. The simplest division of volume in the reconstruction defines each subunit of the pilus as a pair of approximately globular domains. This subunit definition may not, however, correspond to the boundaries of a single PapA molecule. Given the structural variations reported below, the PapA molecules are probably extended along the one-start helix. The degree to which they are extended has not been determined.

The P-pili are subject to damage by shear. The mechanical resilience of P-pili is provided by PapA, the major structural protein. Whereas intact P-pili are straight, rigid structures, they are often observed with sharp bends where they appear to have been damaged (Fig. 3). Thin, thread-like structures are also

FIG. 1 Electron micrographs of negatively stained P-pili and the power spectrum of a single filament. *a*, Electron micrograph of negatively stained pili. *b*, Electron micrograph of a single, isolated filament with its power spectrum. The most prominent layer line is the 7th ($n=1$; $l=7$) at a spacing of 24.9 Å, corresponding to the pitch of the one-start helix. Magnification bar, 500 Å.

METHODS. P-pili were isolated as described² using hyperpilated *E. coli* ClapRHU845 (gift from R. Goldstein) harbouring the plasmid pPAP5. Preparations of purified P-pili were negatively stained with phosphotungstic acid, pH 7. Samples were rotary shadowed with Pt and stabilized with a layer of carbon. Grids were examined and photographed in a Philips CM12 electron microscope at 100 kV. Negatively stained samples were photographed under low-dose conditions, $\sim 5\text{e}^- \text{Å}^{-2}$.



observed, at times spanning the gap between two pilus fragments. Pili that have been rotary-shadowed exhibit more threads, probably due to the greater chemical and mechanical stresses that occur during preparation of these specimens.

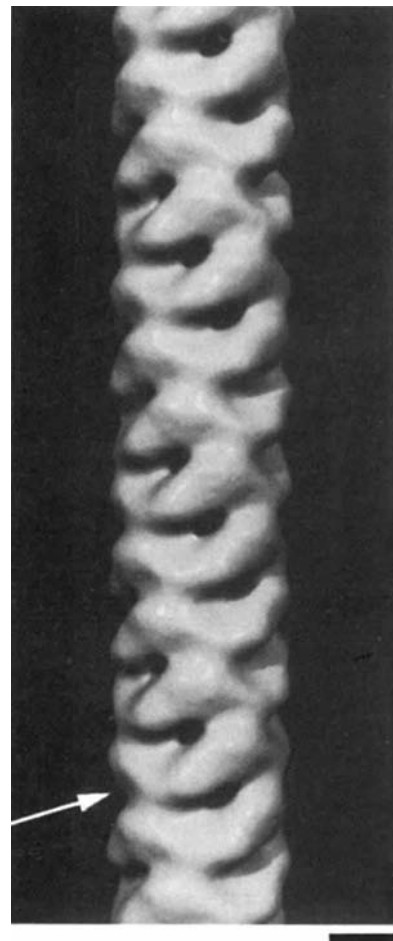
The thread-like structures occasionally appear to have an open helical structure, and look very similar to the tip fibrillae⁶ formed by the minor structural proteins, PapE and PapF. These similarities suggest that the thread-like structures formed by mechanically stressing subunits of PapA have a molecular structure

similar to that of PapE and PapF. If this is correct, the interactions of the pilins with PapD, the chaperone responsible for translocation of the proteins across the periplasm, may be conserved over more than just their C-terminal regions. Proteins that cap a macromolecular helix may have structural motifs similar to those of the major structural protein¹⁴. The similarities among the amino-acid sequences of the pilins^{3,8,9,15} support this possibility for pili.

Plasticity of P-pili allows bacteria to maintain attachment to

FIG. 2 Computer rendering of a three-dimensional reconstruction of P-pili. The contour level was chosen so the volume of each subunit corresponds to that of a single PapA protein. The arrow indicates the direction of the one-start helix. One repeat of the helix contains 23 subunits in 7 turns, with 3.28 subunits per turn. Image generated by the program 'voxray' (by courtesy of J. Griffith; other display programs by C. Owen, T. Tibbitts, C. Potter). Magnification bar, 50 Å.

METHODS. Helical reconstruction programs from the MRC, Cambridge, UK were used with modification¹⁹. The pili chosen for reconstruction were fairly straight. These filaments were further straightened with a cubic splinefit²⁰, and apodized. The helical reconstruction then proceeded as described²¹. Helicity of the filaments was checked, the axial shifts of the pili from the box centre, and the tilt of the pili from the plane perpendicular to the electron beam axis were corrected. The average value for helicity, $Q_{\min}$, was $23.5^\circ \pm 7^\circ$; a filament with $Q_{\min} = 12^\circ$ was chosen as the reference. The reference filament was used to align and scale the pili with respect to each other, and to determine the polarity of the filaments²². To improve the signal-to-noise ratio, near and far data from a filament were averaged and used to determine the relative polarity of each pilus. An average reconstruction was computed, and the individual data sets were refitted to this average. A new average was computed from the 16 averaged near/far sides with self-consistent polarity. Individual data sets were compared to this new reconstruction, with 18 of the 21 filaments being self-consistent. This average reconstruction was then used as the reference for computing a reconstruction with Bessel-filtered layer lines²³. The cycle was repeated twice more to remove bias toward the reference. In the final round of averaging, only filaments with residuals 5° to 26° were included; $Q_{\min} = 13^\circ \pm 5^\circ$. Twenty-seven near/far pili data sets were averaged with respect to the current reference, creating an average reconstruction computed from $\sim 3,500$ pilin subunits.



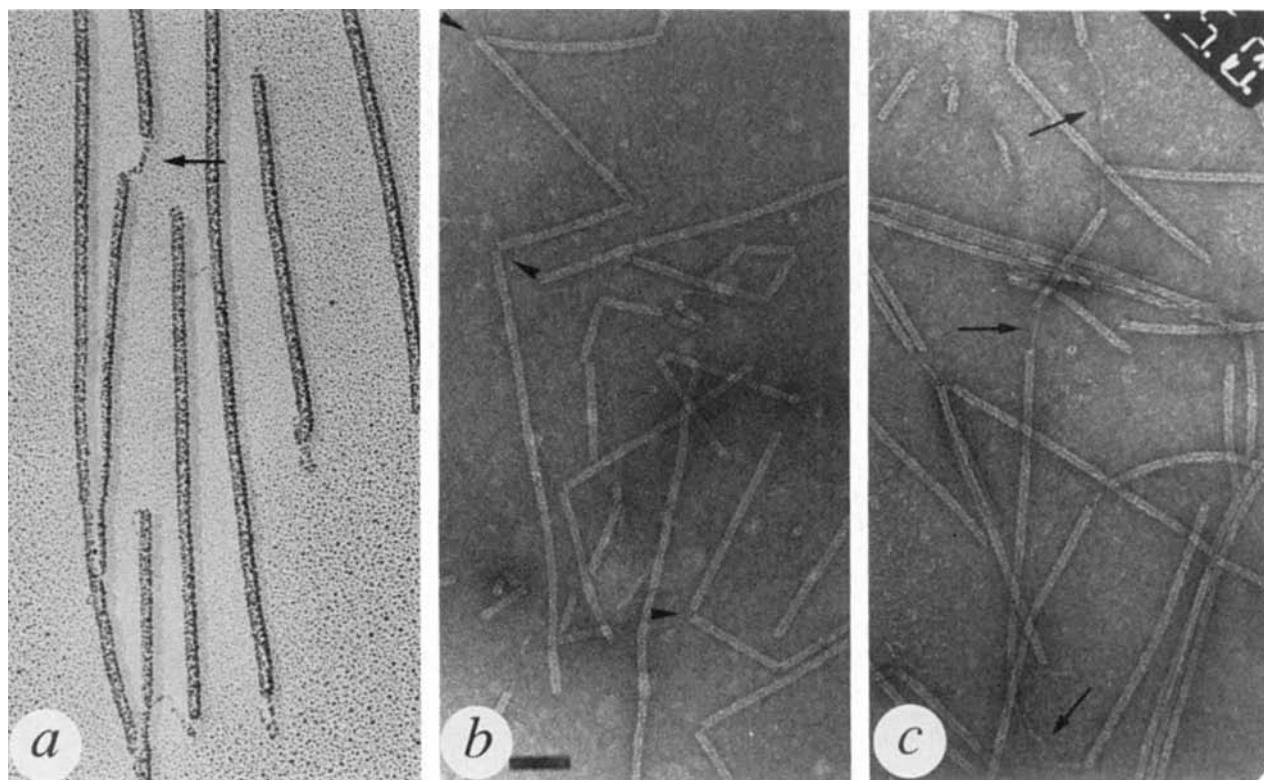


FIG. 3 Electron micrographs of damaged pili containing thread-like structures (arrows) and sharp bends (arrowheads). Mechanically sheared P-pili exhibit a substantial number of these sharp bends, with the pili continuous, but obviously damaged at the turn. Thread-like structures of ~ 20 Å in diameter can often be seen at the ends of pili

or spanning breaks in the pili (compare with chemically treated type 1 pili²⁴). In many cases, these structures appear to have an open, helical structure with a pitch of ~ 150 Å. a, Rotary-shadowed pili; b, c, negatively stained pili. Magnification bar, 500 Å.

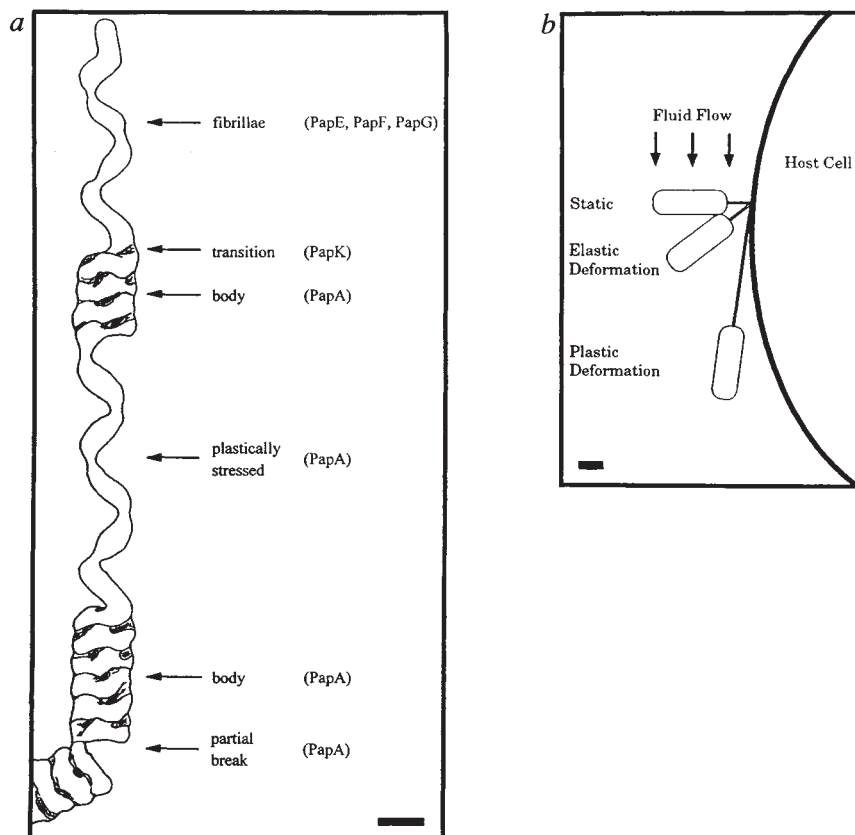


FIG. 4 Models for a, the mechanical unravelling of the pili, and b, the effect that the extensive plasticity has on the attachment of the pili to target cells. b, The maximum permissible movement for an *E. coli* cell which remains attached, assuming either elastic or plastic deformation of the pili. The length of the pilus is increased from 1 μm to 1.15 μm or 5.0 μm , respectively, for elastic versus plastic deformation. Each cell may have dozens of pili involved in attachment to the host cell. Only a single pilus is shown here for simplicity. The fibrillae contain PapE, PapF and PapG⁶. PapE and PapF may have a configuration similar to the extended configuration of PapA. PapA in the body of the pilus has a tightly wound configuration, and makes bonds to PapA molecules on adjacent turns of the one-start helix. Under mechanical stress, the PapA subunit can switch to an extended conformation, and bonds between adjacent turns of the one-start helix break, resulting in an open helical structure. At partial breaks in the pilus structure, bonds between adjacent turns of the one-start helix break, but the conformation of only a single PapA is altered. Magnification bars, 50 Å (a) and 1 μm (b).

host cells under conditions that would otherwise remove them from the target tissue. The transformation of the pilus from a rigid, rod-like structure to a more flexible, open-helical form requires two things to happen; the bonds between PapA on adjacent turns of the one-start helix must break; and there must be a rotation between adjacent domains along the one-start helix. A sketch of this transformation is shown in Fig. 4. The result of these events is the lengthening of the pilus by a factor of about five. This intrinsic plasticity allows the pilus to support tension over a broad range of lengths. This in turn makes it possible for a much larger number of pili to support tension simultaneously, providing stronger adhesion of the bacterium to the target tissue. Support of the bacterium by rigid protein rods would be less advantageous, as a shear field due to urine flow could lead to their failure one at a time.

The mechanical resilience of the pili suggests that PapA is fibrous and highly extended along the one-start helix of the pilus. The combination of these data suggests that all the structural proteins of the pili (in particular, PapA, PapE and PapF) are highly extended fibrous proteins. There are currently no structural data that address this possibility.

Adhesion pili join flagella as a second class of bacterial appendages capable of a purposeful structural transition in response to mechanical stress. The two helical forms of pili are reminiscent of the mechanically induced structural transitions in bacterial flagella. In both bacterial flagella¹⁶⁻¹⁸ and P-pili, the

structural changes in response to stress may be critical to bacterial survival. In the case of pili the structural transition may help maintain bacterial adhesion so that colonization can occur. □

Received 25 May; accepted 18 November 1994.

1. Ørskov, I., Ørskov, F. & Birch-Andersen, A. *Infect. Immun.* **27**, 657–666 (1980).
2. Gong, M. & Makowski, L. *J. molec. Biol.* **228**, 735–742 (1992).
3. Båga, M. et al. *J. Bact.* **157**, 330–333 (1984).
4. Båga, M., Norgren, M. & Normark, S. *Cell* **49**, 241–251 (1987).
5. Lindberg, F., Lund, B., Johansson, L. & Normark, S. *Nature* **328**, 84–87 (1987).
6. Kuehn, M. J., Heuser, J., Normark, S. & Hultgren, S. J. *Nature* **356**, 252–255 (1992).
7. Jacob-Dubuisson, F., Heuser, J., Dodson, K., Normark, S. & Hultgren, S. J. *EMBO J.* **12**, 837–847 (1993).
8. Denich, K. et al. *Infect. & Immun.* **59**, 3849–3858 (1991).
9. Lindberg, F., Lund, B. & Normark, S. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1891–1895 (1986).
10. Kuehn, M. J., Normark, S. & Hultgren, S. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10586–10590 (1991).
11. Slonim, L. N., Pinker, J. S., Bränden, C.-I. & Hultgren, S. J. *EMBO J.* **11**, 4747–4756 (1992).
12. Hultgren, S., Jacob-Dubuisson, F., Jones, C. H. & Bränden, C. I. *Adv. Prot. Chem.* **44**, 99–123 (1993).
13. Brinton, C. C. *Trans. N.Y. Acad. Sci.* **27**, 1003–1054 (1965).
14. Makowski, L. *J. molec. Biol.* **228**, 885–892 (1992).
15. Marklund, B. I. et al. *J. Bact.* **157**, 330–333 (1992).
16. Kamiya, R. & Asakura, S. *J. molec. Biol.* **106**, 167–186 (1976).
17. Macnab, R. M. & Ornston, M. K. *J. molec. Biol.* **112**, 1–30 (1977).
18. Calladine, C. R. *J. molec. Biol.* **118**, 457–479 (1978).
19. Owen, C. & DeRosier, D. J. *J. Cell Biol.* **123**, 337–344 (1993).
20. Egelman, E. *Ultramicroscopy* **19**, 367–374 (1986).
21. DeRosier, D. & Moore, P. J. *J. molec. Biol.* **52**, 355–369 (1970).
22. Amos, L. J. *J. molec. Biol.* **99**, 65–73 (1975).
23. Morgan, D. G. & DeRosier, D. J. *Ultramicroscopy* **46**, 263–285 (1992).
24. Hultgren, S. J., Normark, S. & Abraham, S. N. *A. Rev. Microbiol.* **45**, 383–415 (1991).

ACKNOWLEDGEMENTS. We thank M. F. Gong, S. Hultgren and R. Striker for discussion, and C. Owen, T. Tibbitts and J. Griffith for sharing computer programs and for discussion. Supported by grants from the NIH and NSF.

The crystal structure of bluetongue virus VP7

Jonathan Grimes*, Ajit K. Basak*†, Polly Roy*†‡ & David Stuart*§

* The Laboratory of Molecular Biophysics, University of Oxford, Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK

§ The Oxford Centre for Molecular Sciences, New Chemistry Building, South Parks Road, Oxford OX1 3QT, UK

† NERC Institute of Virology and Environmental Microbiology, Mansfield Road, Oxford OX1 3SR, UK

‡ School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama 35294, USA

BLUETONGUE virus (BTV), a representative of the orbivirus genus of the Reoviridae, is considerably larger (at 80 nm across), and structurally more complex, than any virus for which we have comprehensive structural information. Orbiviruses infect mammalian hosts through insect vectors and cause economically important diseases of domesticated animals¹. They possess a segmented double-stranded RNA genome within a capsid composed of four major types of polypeptide chains¹. An outer layer of VP2 and VP5 is removed as the virus enters the target cell, to leave an intact core within the cell. This core is 70 nm across and composed of 780 copies of VP7 (*M*, 38K) that, as trimers, form 260 'bristly' capsomeres clothing an inner scaffold constructed from VP3 (*M*, 103K)². We report here the crystal structure of VP7 from BTV serotype 10, which reveals a molecular architecture not seen previously in viral structural proteins. Each subunit consists of two domains, one a β -sandwich, the other a bundle of α -helices, and a short carboxy-terminal arm which might tie trimers together during capsid formation. A concentration of methionine residues at the core of the molecule could provide plasticity, relieving structural mismatches during assembly.

The structure determination at 2.6 Å resolution, of 6 non-crystallographically related copies of the polypeptide chain arranged as two independent trimers, is described in Table 1. The present model has a crystallographic *R*-factor of 18.4% with

strict non-crystallographic constraints and r.m.s. deviation from ideal bond lengths of 0.016 Å. All 349 residues of the mature protein are seen in the electron density maps.

Using the published electron micrographic reconstructions, we can unambiguously orientate the trimer within the core; the molecular 3-fold axis is perpendicular to the core surface and the broader base of the trimer contacts an internal network of VP3 molecules (Fig. 1a). The VP7 trimer measures 85 Å along the 3-fold axis and 65 Å in other directions (Fig. 1c). Each subunit consists of two distinct domains (Fig. 1b). The smaller domain forms the outer surface of the core and contains the central $\frac{1}{3}$ of the polypeptide chain, residues 121–249, folded into a β -sandwich. The larger domain (residues 1–120 and 250–349) is composed of α -helices and long extended loops and is less external to the core, interacting with the underlying VP3 and also, side-to-side, with other trimers of VP7. The two domains are twisted with respect to each other around the 3-fold axis of the molecule, so that the top domain of one monomer rests on the lower domain of a related subunit (Fig. 1c). The antiparallel β -sandwich of the upper domain is similar to the jelly roll motif common to most viral capsid proteins³. The topologies of the two β -sheets are A'BIDG and CHEF, which corresponds to a jelly roll with an extra strand clipped onto the inner, BIDG sheet, making it very similar to the jelly roll of influenza virus haemagglutinin⁴ (Fig. 2a); indeed the similarity is comparable to that between the satellite tobacco necrosis virus (STNV) and tomato bushy stunt virus (TBSV) capsid proteins for which a common ancestor has been proposed⁵. The larger domain is composed of 9 helices; 5 in the N-terminal portion and the remainder in the C-terminal region (Fig. 1b). The helices are packed tightly together in a complex pattern. At the heart of this domain is an extensive hydrophobic core formed predominantly of side chains contributed from an antiparallel 4-helix bundle (helices 1, 3, 7 and 8), with the right-handed up-down-down-up topology seen in phospholipase C (ref. 6). However, we cannot detect convincing structural evidence for homology between the lower domain of VP7 and any other protein. Bundles of helices are observed on the interior of the capsid of small insect viruses such as flockhouse virus⁷, but the topology appears quite different.

TABLE 1 Structure determination

Processing statistics									
Data set	Number of crystals	Maximum resolution (Å)	Total reflections	Unique reflections	% complete	R_{merge} (%)	MFID (%)	FM	$\langle F \rangle / \langle E \rangle$
Native	3	2.6	314,113	57,697	81	10.5	—	—	—
AuCN	4	3.0	398,191	44,434	96	8.6	11.34 (4 Å)	0.2 (3.5 Å)	0.7 (3.5 Å)
Refinement statistics									
Resolution range			15 Å–2.6 Å						
Number of reflections			54,361 (negatives excluded)						
R_{factor}			18.4%						
Number of protein atoms			2,703						
Number of other atoms			—						
$\langle \text{Temperature factor} \rangle$ backbone			23.8 Å ²						
r.m.s. deviations from ideality			Bond lengths: 0.016 Å				Bond angles: 1.85°		
Temperature factor deviations			Bond related: 4.0 Å ²						

Full details of the structure determination will be presented elsewhere. In brief: VP7 of BTV, serotype 10, was crystallized, using the hanging drop vapour diffusion technique¹⁹ from 10% w/v PEG 3500 in 0.1 M sodium acetate at pH 4.6 (protein concentration 12 mg ml⁻¹). The crystals belong to space group $P2_1$ with cell dimensions $a = 83.6$ Å, $b = 110.1$ Å, $c = 129.8$ Å, $\beta = 103.1^\circ$. Protease inhibitors added to the protein solution before crystallization prevented proteolysis. Crystals appeared within 3–4 h and reached maximum dimensions of $1.0 \times 0.5 \times 0.35$ mm³ within 2–3 days. The crystallographic asymmetric unit contains two trimers and roughly 45% solvent. Data were collected at 20 °C from both native and derivative crystals using an in-house Marresearch imaging plate system and a Rigaku RU200H rotating anode X-ray source. CuK α radiation was selected with a graphite monochromator and collimated to dimensions of 0.3×0.3 mm². The heavy-atom derivative was made by soaking a crystal in a 3 mM solution of gold cyanide. Data were processed using DENZO and SCALEPACK²⁰. The positions of the 6 heavy-atom sites were determined using GROPAT²¹ and refined using MLPHARE²². The resultant SIR phases were refined by cyclic solvent flattening using GAP (J. M. Grimes and D. I. Stuart, unpublished program) and CCP4 programs²³. 6-fold NCS real-space averaging was then started using GAP, with initial NCS operators calculated from the heavy-atom sites. Envelopes for each trimer were defined using GAP. SIR and averaged phases were combined²⁴. During the cyclic averaging, envelopes were improved and the molecular orientations repeatedly refined. This procedure worked well for trimer 1 but imperfectly for trimer 2. Thus the real space correlation coefficient (defined as $\sum_i \sum_j \rho_{i,j} \times \langle \rho_{i,j} \rangle / \sum_i \sum_j \langle \rho_{i,j} \rangle^2$ where $\rho_{i,j}$ is the electron density for copy i at pixel j in the molecular unit and $\langle \rho_{i,j} \rangle$ is the mean value for the electron density, averaged over the 6 independent values) for trimer 1 rose to 90% but only to 85% for trimer 2. Rigid body refinement of a previously solved fragment of VP7 (A.K.B. et al., manuscript in preparation) against these data using XPLOR²⁵ provided improved NCS operators which improved the real space correlation coefficients to 93.8% and 93.5% from trimers 1 and 2, respectively, and provided an easily interpreted map at 3.5 Å resolution into which a model was quickly built. Coarse-grained phase extension in reciprocal space was done using this model in XPLOR. Refinement was initiated at 3.5 Å, and extended to 2.6 Å (R -factor 28.1%). Averaging and solvent flattening at 2.6 Å, starting from these phases, led to a reciprocal space R -factor of 12.6% and correlation coefficient of 96% and a map of exceptional quality. Building the structure with FRODO²⁶ was trivial. XPLOR refinement using strict NCS constraints, a solvent mask and all data from 15 Å to 2.6 Å led to the current model which has 90% of its residues in the most favoured region of the Ramachandran plot. $R_{\text{merge}} = \sum_i \sum_h |I_{i,h} - \langle I_h \rangle| / \sum_i \sum_h \langle I_h \rangle$ where h are unique reflection indices, $I_{i,h}$ are intensities of symmetry redundant reflections and I_h is the mean intensity. MFID is the mean fractional difference between the native and heavy-atom derivative structure factor amplitudes. FM is the overall figure of merit: the cosine of the mean phase error of the phases. $\langle F \rangle / \langle E \rangle$ is the phasing power: the r.m.s. value of F_H divided by the r.m.s. lack of closure error. $R_{\text{factor}} = \sum_h ||F_o|_h - |F_c||_h| / \sum_h |F_o|_h$ where h defines the unique reflections.

The trimer is a robust building block: the interactions between subunits are extensive, 5,200 Å² of surface area is lost by each subunit upon trimer formation (both domains contribute to the interactions), more than the typical value for an oligomeric protein of this size⁸. However, there are significant cavities at the heart of the trimer, which lie along the 3-fold axis. These cavities have an overall volume of some 3,000 Å³ and are lined predominantly by uncharged residues. Hydrophobic cavities are also seen in certain picornaviruses, imparting essential flexibility to the capsid⁹.

How do the VP7 molecules achieve the combination of structural variation with precision necessary to assemble 260 trimers side-by-side in a closed shell with the observed $T = 13$ symmetry? We believe that the C-terminal residues, from 333 onwards, are crucially involved in the interactions between trimers. Mutagenesis¹⁰ shows that addition of extra residues to the C terminus abolishes core formation. Furthermore these residues are at the greatest distance from the molecular 3-fold axis and form a highly flexible hydrophobic loop and the final helix, suggesting that this part of the structure may be able to adopt a number of conformations. This structure may act as an abbreviated version of the 'arm' which, by conformational switching, aids the formation of the $T = 7$ lattice of simian virus 40 (SV40)¹¹.

To understand assembly fully we must also understand how VP7 clothes the inner network of VP3 molecules. This involves a symmetry mismatch in the interactions between the bases of the 260 trimers of VP7 and a smaller number of VP3 molecules

(60 (ref. 12) and 120 (P. P. C. Mertens and J. N. Burroughs, personal communication) have both been proposed). How is this mismatch ameliorated sufficiently to stabilize the capsid during and after assembly? The base of the VP7 trimer is formed largely by helix 2, and the somewhat flexible loops leading to and from this helix. These elements contrive to form an exceptionally flat surface (Fig. 1d) packed with large hydrophobic side chains. The propensity of this face of the molecule for protein–protein interactions is shown in the crystal where base-to-base packing occludes an area (2,000 Å²) over twice as large as a typical antibody combining site. This surface appears to be tailored to its biological role, achieving tight association by predominantly hydrophobic interactions, with perhaps some leeway to relieve mismatches.

There may also be a further mechanism at work, involving an unusual aspect of the internal structure of the α -helical domain of BTV VP7; internal fluidity may aid capsid closure. Methionine residues have previously been implicated in conformational switching¹³ and VP7 contains over twice as many as would be expected in a protein of this size¹⁴ with a cluster of 5 methionine and 2 cysteine residues at the heart of the lower domain (Fig. 1c). In our crystals of the recombinant protein (grown and maintained in non-reducing conditions), the two cysteine residues (attached to helices 1 and 3) do not form a disulphide bond, in spite of being almost ideally orientated to do so, with the sulphur atoms only 3.6 Å apart. However, the majority of VP7 subunits derived from viral capsids possess this disulphide bond¹⁵. It is possible that on assembly the plasticity of the hydrophobic core

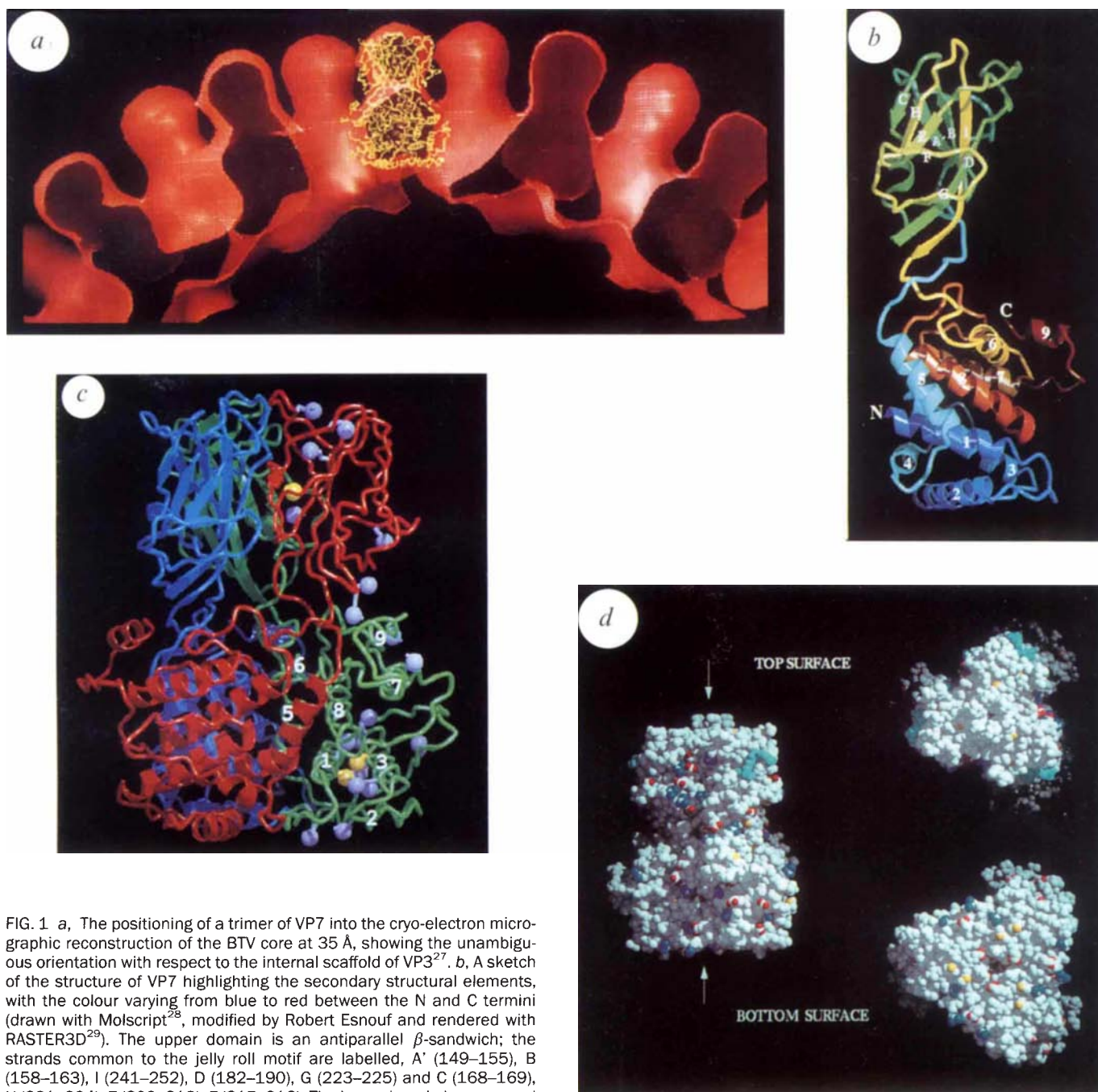


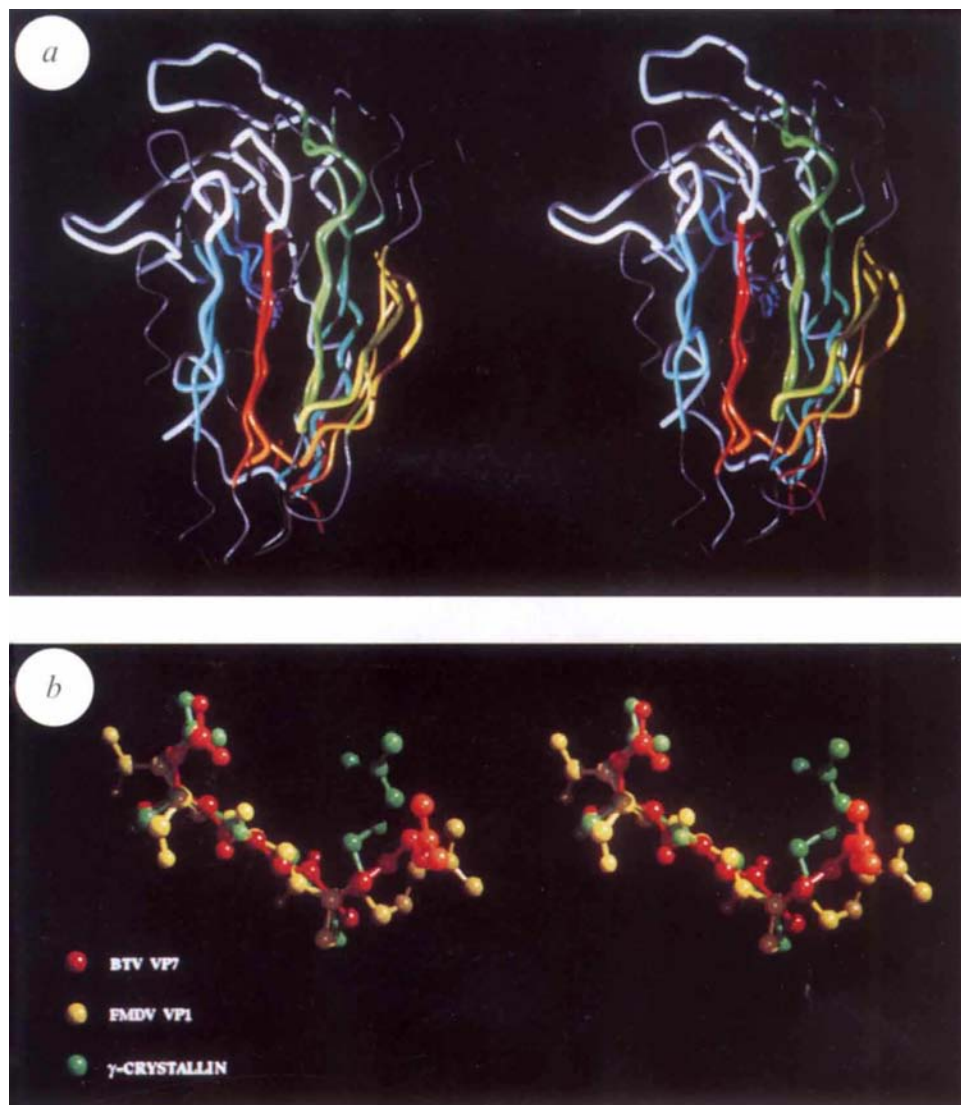
FIG. 1 *a*, The positioning of a trimer of VP7 into the cryo-electron micrographic reconstruction of the BTV core at 35 Å, showing the unambiguous orientation with respect to the internal scaffold of VP3²⁷. *b*, A sketch of the structure of VP7 highlighting the secondary structural elements, with the colour varying from blue to red between the N and C termini (drawn with Molscript²⁸, modified by Robert Esnouf and rendered with RASTER3D²⁹). The upper domain is an antiparallel β -sandwich; the strands common to the jelly roll motif are labelled, A' (149–155), B (158–163), I (241–252), D (182–190), G (223–225) and C (168–169), H (231–234), E (209–212), F (215–218). The lower domain is composed of 9 α -helices and is clearly split into 2 sections, with the C-terminal region packing on top of the N terminus. The helices are labelled sequentially by position in the amino-acid sequence, 1 (1–17), 2 (27–44), 3 (56–74), 4 (89–95), 5 (113–117), 6 (263–271), 7 (276–288), 8 (304–323) and 9 (341–347). The B-factors (which indicate the degree of flexibility) for the main chain atoms rise above 60 Å² for residues 20–21, to about 50 Å² for 145–146 and 238–239 and to above 60 Å² for residues 333–338. *c*, The trimer of BTV VP7, highlighting the large number of methionine (purple) residues in a monomer. Cysteine residues are coloured yellow. For clarity the Ca trace of the bottom domain of the green monomer, looking along the 4-helix bundle, and the top domain of the red monomer, looking along the β -sandwich, are shown as a thin tube²⁸. The top domain interactions are mediated

by 2 extended loops from each monomer, the long D–E top loop (196–208) and the A–A' loop which clips onto the inner face of the β -sandwich of an adjacent subunit. The interactions between the lower domains in the trimer are complex, mostly involving helices 5 and 6 and the associated loops. In addition, helix 9, formed from the C-terminal residues of the protein and lying on the outside of the trimer, packs against residues from the top domain of another subunit. *d*, CPK representation of BTV VP7 looking vertically onto the base and top (right) and perpendicular (left) to the 3-fold, highlighting the large areas of uncharged hydrophobic surfaces. Nitrogens and oxygens of the side chains of charged residues are coloured blue and red, respectively, and sulphurs are coloured yellow²⁸. The tripeptide RGD (residues 168–170) is highlighted in cyan.

between helices 1, 3 and 8 allows some repacking of the secondary structural elements of the lower domain, altering the orientation of helix 2 to relieve the structural mismatches with the lower symmetry scaffold of VP3. This would change the relative positions of the cysteine residues and the resultant structure might

then be locked in place by disulphide bond formation. Evidence from site-directed mutagenesis accords with this hypothesis; although the disulphide bond is not required for the formation of core-like particles¹⁰ its presence further stabilizes the resultant particles (P.R., unpublished data).

FIG. 2 *a*, Comparison of the trace of the top domain of VP7 (134 to 253) with the jelly roll of haemagglutinin. VP7 has the thicker tube trace and those sections of the structure that have been equivalenced are coloured blue to red between the N and C termini²⁸. The superimposition of *Cas*, with SHP³⁰ (using an automatic definition of 'equivalent' residues), found 74 structural equivalences (64% of the top domain of VP7) with r.m.s. deviation between equivalenced atoms of 2.97 Å. This is not dissimilar to the fit of 66% of STNV to TBSV with an r.m.s. deviation of 2.72 Å between equivalenced *Cas*. *b*, Stereo superimposition of the *Cas* of the RGD motifs of FMDV (yellow) and γ -crystallin (green) onto VP7 (red), gave r.m.s. deviations between *Cas* of 0.15 Å and 0.08 Å, respectively²⁸.



The structure of VP7 may throw some light on cell attachment. BTV infects both insect and mammalian cells although the respective cell surface receptors have not been identified. Cores bind to mammalian cells but show very low infectivity¹⁶, perhaps because these cells cannot correctly internalize them. They are, however, highly infectious against cells of the insect vector *Culicoides variipennis*, suggesting a secondary route of infection using a different mechanism of entry. It is therefore interesting that an Arg-Gly-Asp (RGD) tripeptide (168–170) is conserved

in almost all known sequences of orbivirus VP7s despite being situated in an exposed position on the core. Such tripeptides are central to many integrin-dependant cell adhesion processes and indeed the triplet in VP7 is in a conformation similar to that seen in other integrin binding RGDs^{17,18} (Fig. 2*b*). The hypothesis that this sequence is involved in cell attachment is supported by studies that show that the BTV cores compete against foot-and-mouth disease virus for its receptor, an integrin (T. Collen and P. P. C. Mertens, personal communication). □

Received 29 September; accepted 22 November 1994.

- Knudson, E. L. & Monath, T. P. in *Virology* (eds Fields, B. N. & Knipe, D. M.) 1405–1433 (Raven, New York, 1990).
- Roy, R., Marshall, J. J. A. & French, T. J. in *Bluetongue Virus* (eds Roy, P. & Gorman, B. M.) 43–87 (Springer, Heidelberg, 1990).
- Stuart, D. I. *Curr. Opin. struct. Biol.* **3**, 167–174 (1993).
- Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366–373 (1981).
- Rossmann, M. G. et al. *J. molec. Biol.* **165**, 711–736 (1983).
- Hough, E. et al. *Nature* **338**, 357–360 (1989).
- Fisher, A. J. & Johnson, J. E. *Nature* **361**, 176–179 (1993).
- Janin, J., Miller, S. & Chothia, C. *J. molec. Biol.* **204**, 155–164 (1988).
- Oliveira, M. A. et al. *Structure* **1**, 51–68 (1993).
- Le Blois, H. & Roy, P. *J. Virol.* **67**, 353–359 (1993).
- Liddington, R. C. et al. *Nature* **354**, 278–284 (1991).
- Hewat, E. A., Booth, T. F. & Roy, P. *Virology* **189**, 10–20 (1992).
- Georgiadis, M. M. et al. *Science* **257**, 1653–1659 (1992).
- Creighton, T. E. *Proteins* (Freeman, New York, 1993).
- Eaton, B. T. et al. *Virology* **180**, 687–696 (1991).
- Mertens, P. P. C., Burroughs, J. N. & Meilior, P. S. *Infection of Culicoides Vector Species by Different Orbivirus Particle Types* (International Congress of Virology, 1993).
- Logan, D. et al. *Nature* **362**, 566–568 (1993).
- Wistow, G. et al. *J. molec. Biol.* **170**, 175–202 (1983).

- Mcherson, A. J. *Preparation and Analysis of Protein Crystals* (Wiley, New York, 1982).
- Otwinowski, Z. in *Data Collection and Processing* (eds Saywer, L., Isaacs, N. & Bailey, S.) 56–62 (SERC, Daresbury Laboratory, Warrington, UK, 1993).
- Jones, Y. & Stuart, D. I. in *Isomorphous Replacement and Anomalous Scattering* (eds Wolf, W., Evans, P. R. & Leslie, A. G. W.) 39–48 (SERC, Daresbury Laboratory, Warrington, UK, 1991).
- Otwinowski, Z., in *Isomorphous Replacement and Anomalous Scattering* (eds Wolf, W., Evans, P. R. & Leslie, A. G. W.) 80–86 (SERC, Daresbury Laboratory, Warrington, UK, 1991).
- CCP4 *Acta crystallogr.* **D50**, 760–763 (1994).
- Read, R. J. *Acta crystallogr.* **A42**, 140–149 (1986).
- Brunker, A. T. *XPLOR Version 3.0* (Yale Univ. Press, New Haven, 1992).
- Jones, T. A. *Meth. Enzym.* **115**, 157–171 (1985).
- Prasad, B. V. V., Yamaguchi, S. & Roy, P. *J. Virol.* **66**, 2135–2142 (1992).
- Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946–950 (1991).
- Merritt, E. A. & Murphy, M. E. P. *Acta crystallogr.* **D50**, 869–873 (1994).
- Stuart, D. I., Levine, M., Muirhead, H. & Stammers, D. K. *J. molec. Biol.* **134**, 109–142 (1979).

ACKNOWLEDGEMENTS. J.G. and A.K.B. contributed equally to this work. We thank Y. Jones for help and advice; K. Harlos provided data collection support; R. Bryan and R. Esnouf computer support. Thanks also to P. Gouet, P. Mertens and G. Sutton. A.B. is supported by the BBSRC, J.G. by the MRC. The Oxford Centre for Molecular Sciences is supported by the BBSRC and MRC. Coordinates have been deposited with the Protein Data Bank.

Out with the old — in with the new

For your New Year's research resolutions — a new set of nine mouse cytokine ELISAs, a robotic analysis composer, three new centrifuges, a hand held infrared thermometer, and a combine SPM/SEM.

BIOMOLECULAR, Inc., has introduced its new PMC Dataplate 8000 Series Digital **multi-position hotplate/stirrers** (*Reader Service No. 100*). The 8050 and 8090 models allow the user to stir five or nine different samples on the same heating surface. Safety features for the 8000 series include dual sensors in the plate and the probe, temperature limit alarms, heater plate limit setting and heater cutout relay. All parameters are set with scrolling up/down arrow keys. These models feature digital display, PID temperature control, electronic calibration, and selfdiagnostic software. The 8050 and 8090 are available with milled flat aluminium, CERICA solid ceramic, or porcelain-coated stainless steel. These are also available in 100, 115, and 230 VAC with 50/60 Hz.



Biomolecular's five-position hotplate.

For **testing membrane integrity** Sartorius has announced Sartocheck 3. Designed to meet the needs of pharmaceutical production environments, this unit offers six available test programmes (*Reader Service No. 101*). Bubble point, diffusion, multipoint diffusion, and water intrusion are four of the tests featured in this instrument. There is a class 0.1 absolute pressure sensor, individually selectable password protection, and a memory card with a storage capacity of up to 240 test programs. Test documentation is designed to meet the needs of good manufacturing practice and is printed on plain paper with two-colour ink for permanent storage.

For researchers involved with genetic and functional studies of mitochondria, Advanced Biotechnologies, Ltd., has introduced its first **paramagnetic bead for purification of human mitochondria** (*Reader Service No. 102*). Antibodies

raised against whole human mitochondria allow rapid purification of mitochondria directly from cell lysates. It is stated that the usual gradient ultracentrifugation procedure is not necessary and purified mitochondria can be isolated in only 30 min free of extraneous nucleic acids and proteins. Genetic analysis can be performed using the DNA extraction reagents and enzymes included in the kit.

Advanced Separation Technologies has introduced its new Jetstream Column Thermostat, a **column heater/cooler for enhanced HPLC separations** (*Reader Service No. 103*). This unit is said to be particularly useful for chiral separations since lower operating temperatures greatly improve resolution. Linear ramp calculations or steps can be programmed. Selectable parameters include temperature changes during operations and rate of cooling and heating. The temperature range is 0 °C to 80 °C. The column compartment is insulated from the environment, is said to hold four columns with lengths of 3–40 cm and will accommodate precolumns, injector valve and column-switch valve.

The new Cole-Palmer **infrared thermometers** with precision aiming systems feature built-in data logging that allows you to store up to 100 data points (*Reader Service No. 104*). These thermometers feature adjustable emissivity to ensure high accuracy. They calculate minimum, maximum, differential, and average temperature in °C or °F. There is a back-lighting display, a locking trigger for taking unat-



IR thermometer with precision aiming.

tended readings and an audible alarm signals over- and under-range temperature conditions. Four different models are available: the dual-laser model has two laser beams to indicate the diameter of the target area, the crossed-laser model has two laser beams meet at the smallest measurable target area, the single-laser model

pinpoints the exact centre of your target with a bright laser spot, the scope model has an integral optical sighting system to ensure precise target alignment.

A new range of **biotin-conjugated lectins** is now available from Oxford GlycoSystems (*Reader Service No. 105*). Biotinylation is said to allow lectin binding to be detected at high sensitivity without affecting the binding specificity of the lectin. The biotinylated lectins from OGS provide a range of known oligosaccharide binding specificities. They are useful tools for characterizing the oligosaccharides on glycoproteins and other glycoconjugates, locating specific oligosaccharide sequences in tissue sections and as research diagnostics.

Burleigh Instruments, Inc., has introduced piezoelectric actuators to their **micromanipulators** to enhance the performance in difficult electrophysiology applications (*Reader Service No. 106*). The PCS-1000 micromanipulators come as fully assembled, 1- to 3-axis systems designed for high stability and precision with features such as a pivoting headstage mount. The system can generate position steps of tens of microns in 100 microseconds and the Inchworm Motor System can move a micropipette over millimetres with nanometre resolution. The piezoelectric design is intended to maximize step acceleration and to improve the advance through tissues and into cells.

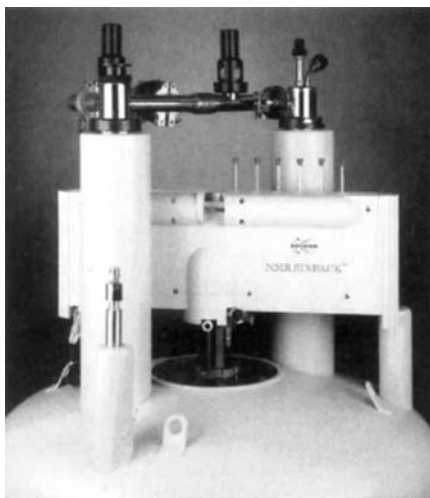
Nine new **mouse cytokine ELISAs** have been added to the Biotrak range of cellular communication assays available from Amersham Life Science (*Reader Service No. 107*). Each assay provides picogram per millilitre sensitivity for its specific cytokine allowing direct measurement in cell culture and serum-plasma samples. Pre-blocked and coated microtitre plates reduce preparation time and results are said to be available in less than 4 h. These new ELISAs are specific for the following mouse cytokines: mIL-2, -3, -4, -5, -6 and -10; mGM-CSF; mIFN- γ and TNF- α . Each kit provides sufficient reagents and controls to prepare one standard curve and to analyse 42 to 44 unknowns in duplicate.

■ Automation

The TECAN **robotic analysis composer** (TRAC) is intended to process a large number of samples for a range of assays including ELISA, ligand binding tests, cell harvesting and cell counting (*Reader*

Service No. 108). The system is designed to meet the high throughput needs of clinical and research laboratories. TRAC uses a robot arm as an automated plate transport device to link Robotic Sample Processors and all associated workstations into one system. The whole system is managed with software developed as an integrated environment for system design, scheduling, control and supervision. Automation capabilities include the ability to link pre-analysis sample preparation with downstream chemistry and analysis. The elements of this component system are the robot arm, liquid dispensing system, microplate gripping hardware, tool parking station, microplate storage racks, plate reader, plate washer, I/O box, personal computer and software. Incubators, plate mixing stations and disposable pipetting tools can be added.

Bruker Instruments, introduces NMR Sixpack, a new compact **NMR sample changer** for its AVANCE, AMX and ARX NMR spectrometers (*Reader Service*



Automatic NMR changer.

No. 109). Designed to improve laboratory efficiency, the Sixpack holds up to six 5 mm or 10 mm samples which the computer changes under full automation. The design of the Sixpack also permits normal insertion of NMR samples, for automated and manual operation.

■ A new spin

Beckman Instruments has announced the new Avanti J-25 **high-performance centrifuge** system (*Reader Service No. 110*). The Avanti J's high-torque drive system accelerates and brakes in times stated as half that of conventional high-speed drives and can deliver 75,600g. Once the centrifuge door has been closed, the Instant Rotor Identification System (IRIS) identifies the rotor automatically for the run eliminating manual entry of rotor information. This centrifuge is imbalance-tolerant to within 5 per cent and tubes can

be easily balanced by eye. A digital display shows set parameters versus actual conditions. Avanti J's vacuum pump creates a one-quarter atmosphere operating pressure in the centrifuge chamber. This reduces the workload of the drive and refrigeration system. It is possible to put a particulate capture filter in line between the vacuum pump and the rotor chamber, allowing the removal of dangerous aerosols and particles that may result from a tube or bottle failure. The CFC-free refrigeration allows all rotors to be run below 4 °C at top speed.



■ **Avanti J-25 centrifuge by Beckman.**

Jouan, Inc. has announced the availability of their 422 series microprocessor-controlled **benchtop centrifuges** (*Reader Service No. 111*). There is the CR422 refrigerated model which allows temperature control from -8 °C to 40 °C, and the CT422 thermostated model, which is capable of controlling chamber temperature from -8 °C to 50 °C. These units provide up to 4200g and have a capacity of 3 L. The rotor insert system is capable of handling special cell culture vessels, such as 250-ml conical tissue bottles — and by changing this insert, tubes and bottles ranging from 0.5 ml to 750 ml. The microprocessor allows the user to programme speed, time, temperature deviation, acceleration and braking rates to define runs, which may be stored in a 50-run memory. Additional features include a control panel that is angled for improved visibility, automatic opening lid, steel guard barrier ring, imbalance detection and dual lid interlock. Sealed bucket with transparent covers permit safe centrifugation of toxic and pathogenic samples.

Ideally suited for quick spins at any workstation is the Sun BIOScience **per-**

sonal microcentrifuge (*Reader Service No. 112*). This lightweight microcentrifuge offers full speed in 3 seconds with an operating force of 14,000g. It has a footprint of 42 in² and is designed for rugged operation with a painted, carbon steel housing to resist corrosion. Available in three different capacities including six 1.5 ml, twelve 1.5 ml and twenty-four 0.5 ml. There is a choice of fixed angle rotor to suit specific needs. Intended applications include nucleic acid preparation, protein separation, PCR product sedimentation, binding studies and a variety of microchemical determinations.

■ Cell tissue culture

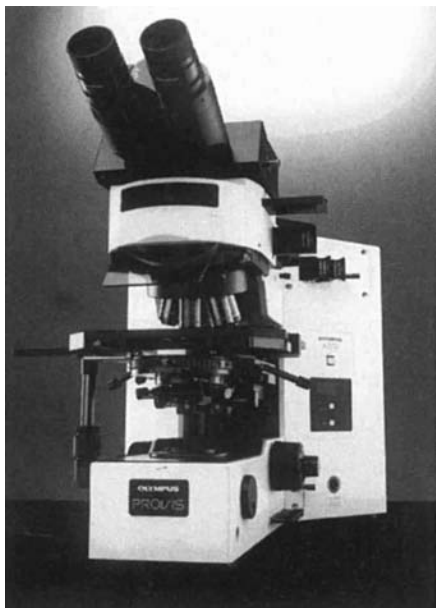
Protein Polymer Technologies has introduced ProNectin F **activated culture ware**, which provides anchorage dependent cells with a surface immediately able to support receptor-mediated cell adhesion, cell spreading and cell growth initiation (*Reader Service No. 113*). The RGD cell attachment ligand of fibronectin is presented directly on the culture surface in the form of ProNectin F, a genetically engineered cell attachment factor. The company states that a wide variety of attachment dependent cells have exhibited faster growth, differentiated functions, and excellent morphology when grown on ProNectin F. The product is sterile, stable and comes in a ready-to-use dish or multi-well plates.

Sigma Cell Culture has introduced SPIT and SPITE medium supplements to its line of **medium supplements** (*Reader Service No. 114*). They contain sodium selenite, sodium pyruvate, insulin and transferrin. SPITE also contains ethanolamine. Both products are for suspension cell cultures and are designed for media that do not contain sodium pyruvate. They are sterile, endotoxin-tested and cell culture-tested. These growth-promoting chemicals are intended to allow cultures to be grown using lower concentrations of fetal bovine serum. Under reduced-serum conditions SPIT has successfully sustained growth of VERO cells and SPITE has contributed to the growth of mouse myeloma, hybridoma, and CHO and hamster kidney cell lines.

Clonetics Corporation has announced EndoPack-CA and MyoPack CASMC cell systems (*Reader Service No. 115*). These cell systems provide normal **human coronary artery endothelial and smooth muscle cells** in optimized low serum (5%) growth medium and all the necessary optimized subculture reagents. This product is intended for research work in all facets of cardiovascular research. All cells are performance tested and test negative for HIV-1, hepatitis-B, mycoplasma, bacteria, yeast and fungi. A certificate of analysis is provided.

■ Microscopy

The PROVIS AX70 universal **research microscope** from Olympus is designed for easy use in all modes of microscopy (*Reader Service No. 116*). The modular design of this instrument offers a choice of configurations and varying degrees of automation. The coaxial coarse and 1 micron fine focusing knobs are positioned low on the AX70 stand, which allows forearms, wrists and hand to rest comfortably on the table surface. The stability and extreme rigidity of this stand, make an excellent platform for photo and video sys-



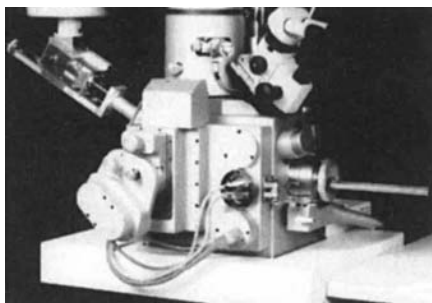
Olympus Provis AX70 Microscope

tems. This microscope is designed to produce images of extremely high resolution and high contrast from its super-wide Universal Plan Fluorite and Plan Apochromat U.I.S. objectives and eyepieces. There is a 12V/100W halogen source for transmitted-light illumination, and for incident light fluorescence there are 100W mercury and 75W xenon sources. Easy access to the optical path allows for the attachment special illuminators. Control of motor-drive units is provided either by a simple controller or a Multi-Controller with RS232C or GPIB interfaces, allowing computer control of the AX70 and its external devices.

The Observer **combined SPM/SEM** is now available on a broad range of scanning electron microscope (*Reader Service No.117*). J K Instruments has announced an exclusive agreement with Peak Instruments to design, manufacture and install interfaces to integrate the TopoMetrix Observer SEM into a broad range of scanning electron microscopes. This system combines the SPM's ability to provide three-dimensional quantitative measurements of surface features down to the atomic scale with the SEM's ability to

offer high scan rates, a larger field of view, a greater depth of field and a spatial resolution of a fraction of a micron. It offers a range of SPM modes, including scanning tunnelling microscopy and atomic force microscopy. Direct vertical measurement by Observer means there is no need to cross-section the sample with SEM.

EnVac is a newly released **scanning electron microscope** from Gresham-CamScan (*Reader Service No. 118*). This instrument enables forensic scientist, archaeologists and other antiquity investigators to use electron microscopy on samples for which normal preparation techniques are unacceptable. Specimens containing moisture can now be examined without elaborate dehydration procedures such as freeze-drying or cryo-preparation followed by coating. EnVac uses a vacuum



EnVac Scanning Electron Microscope.

pumping system which establishes different pressures within the microscope to overcome problems associated with conventional systems. The sample chamber is automatically maintained at a higher pressure than the rest of the electron column, eliminating the need for continual adjustment by the operator. EnVac can be switched easily between low vacuum and high vacuum modes without degrading performance.

■ Labware

The Chemware **teflon PFA and PTFE beakers** from Norton Performance Plastics will withstand heat to 500 °C and 550 °C, respectively (*Reader Service No. 119*). The PFA beaker should be of use for trace metal analysis as this resin is stated to have the lowest concentration of metal ions of all fluoropolymers. It can also be used in cell culture and biotechnology applications where complications can arise from protein binding and adsorption. It is also suggested for digestion and evaporation processes, pilot-scale chip processing with hot acid etchants, catalytic reaction processes and microwave heating. The PTFE beaker is ideal for handling highly aggressive acids that attack glassware and stainless steel. The Chemware beakers are said to be virtually unbreakable and will not shatter as a result of temperature extremes or chemical reactions. These are

non-stick surfaces and have no seams to split or collect contaminants.

A tamper-evident **resealable preservation pouch** that combines the security of a barrier pouch with the ease of a resealable zipper has been introduced by Protective Packaging (*Reader Service No. 120*). Vapor-Loc offers products the protection of a low moisture vapour transmission rate enabling long-term storage, with a top seal that instantly reveals tampering to ensure product integrity. All pouches are custom made and are available in several material combinations including polyester/poly/foil/poly, saran coated polyester/poly and metalized polyester/poly. All bags can be printed on both sides in up to six colours.

Gilson Medical Electronics has introduced two Microman positive-displacement pipettes (*Reader Service No. 121*). The Microman Classic (3 ml–250 ml) is recommended for pipetting of problem liquids such as serum, whole blood, syrup and oil. The Microman Bio (0.5 ml–100 ml) is recommended for PCR and biological samples when the risk of aerosol contamination or carryover must absolutely be avoided. These pipettes are permanently calibrated with unbreakable plastic capillaries and no-touch disposal of capillaries and pistons.

■ Instrumentation

Spectra-Tech has introduced their new complete line of **FT-IR fibre optics** (*Reader Service No. 122*). These new infrared fibre optics offer the ability to analyse samples at remote locations. They are also ideal for analysing samples that are not the optimum shape or size to fit in FT-IR sample compartments. This offers flexi-



Spectra-Tech FT-IR fibre optics.

bility for applications involving hazardous environments, process and reaction monitoring, restricted access sampling vessels, non-destructive sampling, caustic and corrosive environments, and high/low temperature and pressure environments. There are fibre optic sampling probes for remote analysis of liquids, powdered solids, sheet material, reflective solids, microsamples and gases. All accessories connect to the spectrometer through the FiberLink-High Efficiency Transfer Optic, which can be used with most FT-IR spectrometers. These cables transmit IR in the 4000 to 900 cm^{-1} range. Spectra-Tech has choices of fibre optic sampling probes to suit many application.

Dynatech has replaced its MR 5000 microplate reader with its new **MRX microplate reader** (Reader Service No. 123). A completely redesigned instrument for analytical performance, the MRX has UV optics, on-board data reduction software, and integrated self-diagnostics. The instrument does its self-diagnostics check during power-up and before every read, which can be printed out at user-defined



Dynatech MRX microplate reader

intervals for permanent record. The 12 channel optics with a constant reference channel are employed giving a dynamic range of 0.1 to 4.0 OD from 340 nm to 850 nm. Additionally, a verification microplate made from referenced optical materials is available to periodically check and validate optical and mechanical performance. This enables MRX to fulfil GLP and laboratory accreditation requirements. On-board Endpoint Software is standard with MRX offering up to 12 assays per plate, output of raw OD data, threshold, difference and range matrices, a choice of 10 curve-fitting routines, data storage for 100 assays, 20 curve fits and 100 plates of data. Other options include additional filters, barcode, temperature controlled hot plate and Agglutination software.

Fison Instruments has extended the application of its continuous flow isotopic analysis methods to include the **isotopic**

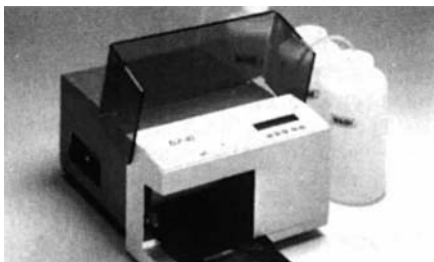


Fison's isotopic sulphur analysis

analysis of sulphur — particularly applicable to geochemical and environmental analysis (Reader Service No. 124). The technique of interfacing an elemental analyser to isotope ratio mass spectrometer offers the dual benefits of high sample throughput with elimination of labour-intensive sample preparations. This technique has now been extended to include isotopic sulphur analysis using Fison's ISOCHROM-EA continuous flow isotope ratio mass spectrometer interfaced with their NA 1500 NCS Elemental Analyser. Applications include food beverage adulteration, soil science, and plant physiology studies.

Miscellaneous

The new Bio-Tek ELP-40 **strip/plate washer** is intended for easy immunoassay washing in diagnostics labs and flexible use in research labs (Reader Service No. 125). Additional functions include the



Bio-Tek ELP-40 strip/plate washer.

adding of reagents and the removal of solutions in cytotoxicity assays. Wash protocols can be customized and the instrument will store up to 75 protocols. Bio-Tek states that the microplate can be shaken, that there is very low residual wash buffer (2 ml) and that it has a footprint of 0.25 m^2 .

Qiagen has introduced the QIAexpress system and Ni-NTA resins for **one-step protein purification** (Reader Service No. 126). Proteins carrying an affinity tag of six consecutive histidine residues can be purified up to 95% homogeneity in just one step, under either native or denaturing conditions. Ni-NTA binds tagged proteins more tightly than other metal-chelating resins due to its unique metal-binding



Techware: from benchtop to bookshelf.

characteristics, leading to more efficient protein purification. Ni-NTA is available as Ni-NTA Agarose for larger scale purification and Ni-NTA Silica for minipreps.

A new **laboratory equipment and supplies catalogue** is now available from Sima-Aldrich Techware (Reader Service No. 127). Intended for use by biochemist, organic chemists, and other life science research professionals, the 1995-96 catalogue features over 10,000 top-quality laboratory equipment and supply product listings. This comprehensive catalogue is divided into 20 different product categories. □

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal

ADVERTISEMENTS

DNA SEQUENCING

Using an ABI 373
Automated Sequencing
System

£45 per run

- PCR products
- M13 inserts
- plasmid inserts
- single stranded DNA
- double stranded DNA
- 300-400 bases of sequence

Please contact us for more information

Molecular Medicine Unit

King's College School of Medicine & Dentistry
Tel: 071-346 3126 Fax: 071-733 3877

Reader Service No.11

PEPTIDE and DNA

Peptide and DNA engineering specialist:

Phosphoryl, Thiophosphoryl, Biotinyl,
Chloromethyl, Cyclic, Dimeric, Partial protected,
FITC, DNP, pNA, Phosphothionate and more ...

IonSpray MS and LC/MS

Synthetic or natural compounds analysis:

Organic, inorganic, or coordination compounds.
Drug metabolites, Enzyme digests, etc.
\$30/sample (MS), \$250/sample (LC/MS)

Call for more information (800-597-7873) or
Fax us sequence for quotation (510-371-1156).

Peptidogenic Research & Co. Livermore, CA

Reader Service No.31